

Using A Natural Experiment to Assess the Effect of Spatial Barriers on Health Service Utilisation

A thesis
submitted in fulfilment
of the requirements for the Degree
of
Master of Science
by
Jayden MacRae



University of Canterbury
2014

Table of Contents

Acknowledgments	vi
Abstract	vi
Chapter 1: Introduction	1
1.1 Health Reforms and Accessibility	1
1.2 A Natural Experiment in the MidCentral District	4
1.3 Research Objectives	8
1.4 Ethics Approval	8
1.5 Thesis Structure	9
Chapter 2: Accessibility of Health Services	10
2.1 The Organisation of New Zealand Health Delivery	10
2.2 Factors that Affect Access	11
2.2.1 Aspatial Factors	12
2.2.2 Spatial Factors	16
2.3 Spatial Accessibility	20
2.4 Existing Literature	22
2.4.1 Potential versus Realised Access	23
2.4.2 The Nearest Service Fallacy	23
2.4.3 Ecological Studies	24
2.4.4 Heterogeneous Populations	25
2.4.5 Geographic Generalisations	26
2.4.6 Characteristics of Rural Versus Urban Meshblock Cal- culations	27
2.4.7 Euclidean Travel Times	27
2.4.8 Road Network Travel Times	29
2.5 Existing Road Network Models	31
2.5.1 New Zealand Models	31

2.5.2	International Models	33
Chapter 3:	Utilisation Metrics in Health Care	36
3.1	Utilisation as a Metric	38
3.2	Emergency Department Attendance	38
3.2.1	Attendance	38
3.2.2	Presentation Urgency and Triage Scores	40
3.3	Hospital Admissions	40
3.3.1	Length of Stay	41
3.3.2	Case Complexity	42
3.4	Outpatient Appointments	43
3.4.1	Attendance	43
3.4.2	Non-attendance	44
3.5	General Practice Utilisation	44
Chapter 4:	Method: Data Quality	46
4.1	Tools and Standards	46
4.2	Last-Best-Known Geocoding	48
4.2.1	Last-Best-Known Process	51
4.3	Corrective Geocoding	51
4.3.1	Data Conversion	53
4.3.2	Loading and Formatting Address Data	55
4.3.3	Address Standardisation	56
4.3.4	Matching	59
4.3.5	Geocoding	60
4.3.6	Validation	61
4.4	Developing a Travel Time Model	62
4.4.1	Existing Attributes	62
4.4.2	Deriving Journey Segments	64
4.4.3	Deriving A Sinuosity Index	64
4.4.4	Deriving Weighted Mean Sinuosity Index	65
4.4.5	Deriving Intersection Density	65
4.4.6	Deriving Elevation Data	66
4.4.7	Deriving Maximum and Total Turning Angles	67

4.4.8	Modeling	68
4.4.9	Model Validation	71
4.5	Calculating Travel Times	73
4.5.1	Road and Routing Network	73
4.5.2	Least Cost Path Analysis	73
4.5.3	Gorge Closure Simulation	76
4.5.4	Final Calculations	77
Chapter 5:	Method: Patient Cohort and Accessibility Measures	78
5.1	Identifying the Patient Cohort	78
5.2	Potential Accessibility	80
5.2.1	Provider-to-Population Ratio	80
5.2.2	Catchments	81
5.2.3	Two-Step Floating Catchment Areas	81
5.2.4	Two-Step Floating Catchments with Distance Decay	82
5.2.5	Patient Choice	83
5.3	Realised Accessibility	84
5.3.1	Utilisation Rate Procedure	84
5.3.2	ED Attendance	85
5.3.3	Hospital Admissions	86
5.3.4	Hospital Outpatients Attendances	86
5.3.5	General Practice Utilisation	86
Chapter 6:	Results	87
6.1	Last-Best-Known Geocoding	87
6.2	Corrective Geocoding	88
6.3	Developing a Travel Time Model	89
6.3.1	Derived Measures	89
6.3.2	Speed & Attribute Correlation	92
6.3.3	Linear Regression	94
6.3.4	Model Comparisons	94
6.3.5	Empirical Validation	97
6.4	Cohort Demography	100

6.5	Potential Accessibility	108
6.5.1	Travel Time Isochrones	108
6.5.2	Provider-to-Population Ratio	108
6.5.3	Catchments	111
6.5.4	Two-Step-Floating Catchments	113
6.5.5	Two-Step-Floating Catchments With Distance Decay .	116
6.6	Realised Accessibility	119
6.6.1	Utilisation by Demographics	119
6.6.2	Choice	121
6.6.3	Intervention Analysis	126
6.6.4	ED Attendance	131
6.6.5	Outpatient Attendance	132
6.6.6	Hospital Admissions	134
6.6.7	GP Attendance	135
Chapter 7: Discussion		137
7.1	Free and Open Source Software and The Scientific Method . .	138
7.2	Geocoding	140
7.2.1	Last-Best-Known Geocode	140
7.2.2	Corrective Geocoding	141
7.3	Travel Time Models	144
7.3.1	Intersection Density	147
7.3.2	Elevation	149
7.3.3	Maximum and Total Turning Angle	150
7.3.4	Linear Model	152
7.3.5	Rule-Based Travel Time	153
7.3.6	Model Comparisons	154
7.3.7	Model Validation	157
7.4	Cohort Demographics and Selection	162
7.5	Potential Accessibility	165
7.5.1	Travel Time Isochrones	166
7.5.2	Provider-to-Population Ratio	167
7.5.3	Catchments	172
7.5.4	Two-Step-Floating Catchments	176

7.5.5	Two-Step-Floating Catchments with Distance Decay	178
7.5.6	Summary of Potential Access	181
7.6	Realised Accessibility	182
7.6.1	Choice	185
7.6.2	ED Attendances	189
7.6.3	Outpatient Attendances	191
7.6.4	Hospital Admissions	195
7.6.5	GP Attendances	198
7.7	Study Limitations	200
Chapter 8:	Conclusion	203
8.1	Implications for MidCentral	206
8.2	Implications for New Zealand	207
8.3	Contributions to the Field	208
Appendices		229
Appendix A:	Interpolation Methodology	230
Appendix B:	Algorithm Implementation Details	233

Acknowledgments

Bernard of Chartres used to say that we are like dwarfs on the shoulders of giants, so that we can see more than they, and things at a greater distance, not by virtue of any sharpness of sight on our part, or any physical distinction, but because we are carried high and raised up by their giant size. - Johannes Parvus (1159)

To my supervisors Prof Simon Kingham and Mr Edward Griffin, thank you. Thanks for your hours of listening and early morning meetings. Thanks for your advice but also your willingness to give me freedom to explore my ideas.

Thank you to Dr Richard Fong and MidCentral District Health Board for supporting this research and for making hospital data sets available.

To my colleague and friend Dr Greg Martin, thank you for listening, for your encouragement, insight and support.

Thanks to my colleagues at Compass Health, for their interest in my area of study and indulging my enthusiasm.

To Compass Health for providing me with the professional development funding to pay for my chosen course of study, thank you.

Thank you to the community of contributors to open source software and data sets, without you sharing your expertise this project in its current form would not have been possible, and my learning would have been much less rich.

Finally, to Caroline for understanding and putting up for two years of my being locked in a study in front a computer in my free time and also for proof-reading my manuscript; thank you.

Abstract

The closure of the Manawatu Gorge in August 2011 caused a change in the travel time for patients living in the eastern area of the MidCentral Health District to their main hospital and health services located in Palmerston North. This presented an opportunity to study the effect a change in travel time and spatial access had on a population before and after such an event. This study used a retrospective cohort design, using routinely collected data from general practice, emergency department, hospital admissions and outpatient services. The investigation was completed using novel geospatial information systems methods to produce high fidelity data for analysis with free and open source software by developing and validating two new methods of improving geocoding data quality and a new travel time prediction model. Potential and realised spatial accessibility measures were calculated for 101,456 patients over 3.5 years while the gorge was both open and closed. Catchment sensitivity analysis and two-step floating catchment area using distance decays presented complimentary evidence of accessibility change during the Manawatu Gorge closure. Analysis of utilisation measures in both primary and secondary care were correlated with travel time. Utilisation of general practice services appeared to be negatively impacted by increased travel time when comparing realised accessibility in a control and intervention group during the gorge closure. It appeared as though other factors affected access to health services to a greater degree than an increase of up to fifteen minutes travel time.

Chapter I

Introduction

1.1 Health Reforms and Accessibility

Health systems have been reforming globally over the past few decades (Saltman and Figueras, 1998). Increasing demand for services has been driven by changes in the age profiles of populations (Hector et al., 2012; Hopman et al., 2009; Mendelson and Schwartz, 1993; Johnson et al., 2013; Suhrchke et al., 2006), sedentary lifestyles and highly refined diets (Goetzel et al., 1998). These in turn contribute to the growing epidemic of chronic disease (Manson et al., 2004). Simultaneously there is a declining resource pool of the health workforce (Buerhaus et al., 2000; Moutier et al., 2011) and an increasing burden of economic pressures on states. Reforms are attempting to address this imbalance by improving the efficiency and productivity of the existing health resources.

Addressing equity of health care to populations has been one target of these reforms. (Hefford et al., 2005; Ham, 1997). Ensuring individuals from all socioeconomic backgrounds can access health services is key to reducing health disparities (Andrulis, 1998). Those with reduced access to primary health care suffer higher rates of avoidable hospitalisations, also known as ambulatory sensitive hospital (ASH) admissions (Rosano et al., 2013; Rizza et al., 2007; Bindman et al., 1995; Parchman and Culler, 1994). An admission is considered to be avoidable when it is caused by a disease that could be managed in primary care or through public health to a level that would avert hospitalisation (Sheerin et al., 2006). These include diseases such as asthma, cellulitis and hyperten-

sion (Purdy et al., 2009; Sheerin et al., 2006; Barnett and Malcolm, 2010).

Secondary health services are those which are provided for specialist care for those with conditions that require more intensive management than general practice can provide. These services are provided through different mechanisms; unplanned presentations requiring urgent attention use Emergency Departments; in-patient admissions exist for those that need extended on-site treatment; out-patient clinics provide access to specialist care on an appointment basis. Secondary care is expensive compared to primary care. One of the aims of good primary care is to keep people from requiring the more expensive and complex secondary based care (Guagliardo, 2004).

Ensuring easy access to specialist in-patient and out-patient services is also a consideration in health care reform. Those who already have chronic disease must manage their condition in conjunction with their health providers. At times patients suffer an exacerbation of their condition even when well managed in primary care. During significant episodes the needs of patients may be beyond what primary care services can provide. Having the ability to access specialist services in such times leads to shorter stays in secondary care facilities and better long term outcomes.

Removing the difficulties that impede access to both primary and secondary health care is one mechanism that is postulated to achieve universal equity. Better access to services leads to better management of health conditions (Teach et al., 2006). There are a variety of reasons that contribute to patients finding it difficult to access health services when they need them.

In some health systems including in New Zealand, accessing primary care comes at a direct cost to patients. Such costs can delay or prevent patients from seeking help in a primary care setting (Andrulis, 1998). These effects can be heightened when emergency and secondary care is provided free-of-charge (Masso et al., 2007; Siminski et al., 2008).

Patients who are financially burdened may opt to delay care or seek it outside of primary care to avoid the additional financial burden it may represent. In some instances the immediate cost may not be the true financial barrier. Where patients have incurred significant debt, they may be discouraged from attending a practice or clinic for fear of the questions, accusations and confrontation that could arise because of it. This behaviour may be exhibited in family groups, where parents refrain from taking their children to a primary care practice because of their own debt.

Having significant travel time or distance to a health service acts as a barrier to service use (Brabyn and Barnett, 2004). Some primary care interactions can be completed adequately over the phone, such as repeat prescriptions. Others often require physical examinations which compel patients to be on-site with the clinician. A patient's ability to physically attend a clinic is a function of their proximity to the facility, their means of transport to it and their mobility. The two highest users of health services, the very old and the very young also happen to be the people that have the least mobility and transport options available to them. Particularly for the elderly, proximity to health services becomes a significant factor in deciding where to reside in retirement (Hodge, 2008).

Current government policy in New Zealand primary care is focusing on integrated family health centres (IFHCs) (Letford and Ashton, 2010; Gauld, 2012). These are intended to improve the co-ordination of care to patients by providing a more multidisciplinary approach (Cumming, 2011). For patients to access multidisciplinary services previously they would usually be required to attend many different physical locations at different times. By co-locating health services and providing some coordination of appointments they have the potential to reduce the travel burden to patients. Integrated Family Health centres are a part of an increasing trend for primary health care providers to consolidate services. Traditional one or two doctor general practices are becoming increasingly less viable under current government policy.

The number of locations from which general practice services are provided is decreasing over time as general practices consolidate services and the number of doctors in the primary care workforce remains constant or decreases. The associated travel burden for patients to those services will therefore be changing. Any change that makes accessing services in primary care less likely would generally be considered counter-productive to existing reform. This follows decades after large scale health reforms in New Zealand where hospital services have been consolidated out of towns into main centres (Barnett, 2000) creating a similar travel burdens to secondary services.

The analysis of spatial accessibility of services is topical given the nature of current and proposed consolidation. Those most likely to be affected are those who live in rural areas because of the way service consolidation naturally gravitates towards main population centres. Spatial barriers to health care have been studied for decades (Joseph and Bantock, 1982; Cromley and Shannon, 1986; Guagliardo, 2004). The literature identifies a consistent theme that the further people are from health services, the less likely they will be to use them (Nemet and Bailey, 2000; Hiscock et al., 2008) and the greater negative impact on their health there is (Jones et al., 1999).

Spatial analysis can provide insight for health policy makers. It can provide the ability to plan and model service location scenarios, as well as monitor use of services by spatial indices.

1.2 A Natural Experiment in the MidCentral District

New Zealand funds health services through twenty-one district health boards (DHBs). They are elected to govern the delivery of health services and expenditure across their designated populations. The populations they serve are all those who reside within well defined geographical borders for each district.

The MidCentral District is located in the lower North Island of New

Zealand, and spans an area of approximately $8,850 \text{ km}^2$. It stretches from the west coast to the east coast and can be seen in Figure 1.1 with cities, towns and their respective populations. The Ruahine and Tararua Ranges form a geographic barrier running from the south-east to north-west dividing the district into natural western and eastern areas. The area of the ranges are clearly demarcated by the absence of roads.

The main route between the two areas is referred to as the Manawatu Gorge, a water gap formed by the Manawatu River. State Highway 3 runs to the south of the river with a rail track running on the northern aspect. Because of the gap formed by the river, the gorge road is relatively flat and provides an almost direct passage between Woodville in the east to Ashhurst in the west. The steep sides of the gorge make it prone to rock-falls and slips.

On the 19th August 2011 the Manawatu Gorge was the site of a slip of epic proportions (Figure 1.2). The road was immediately closed limiting access between the eastern and western areas of the district. There are only two viable alternative road routes across the Tararua ranges; the Saddle Road just to the north of the gorge; and the Pahiatua Track to the South. The road was partially reopened to one lane on the 29th August 2012, and to two-way traffic on the 19th September 2012.

Each of the alternative routes are known to take longer than the relatively flat and direct State Highway 3 route. Palmerston North is the main urban area of the MidCentral District and has the district's public hospital. This resulted in those living in the east bearing an additional burden to attend health services at the hospital. Patients that used general practices services on the other side of the gorge to which they lived also suffered a similar burden.

A rare opportunity to investigate the impact of increased spatial barriers on those accessing health services has presented itself through these circumstances. The gorge was closed for a year. This is enough time

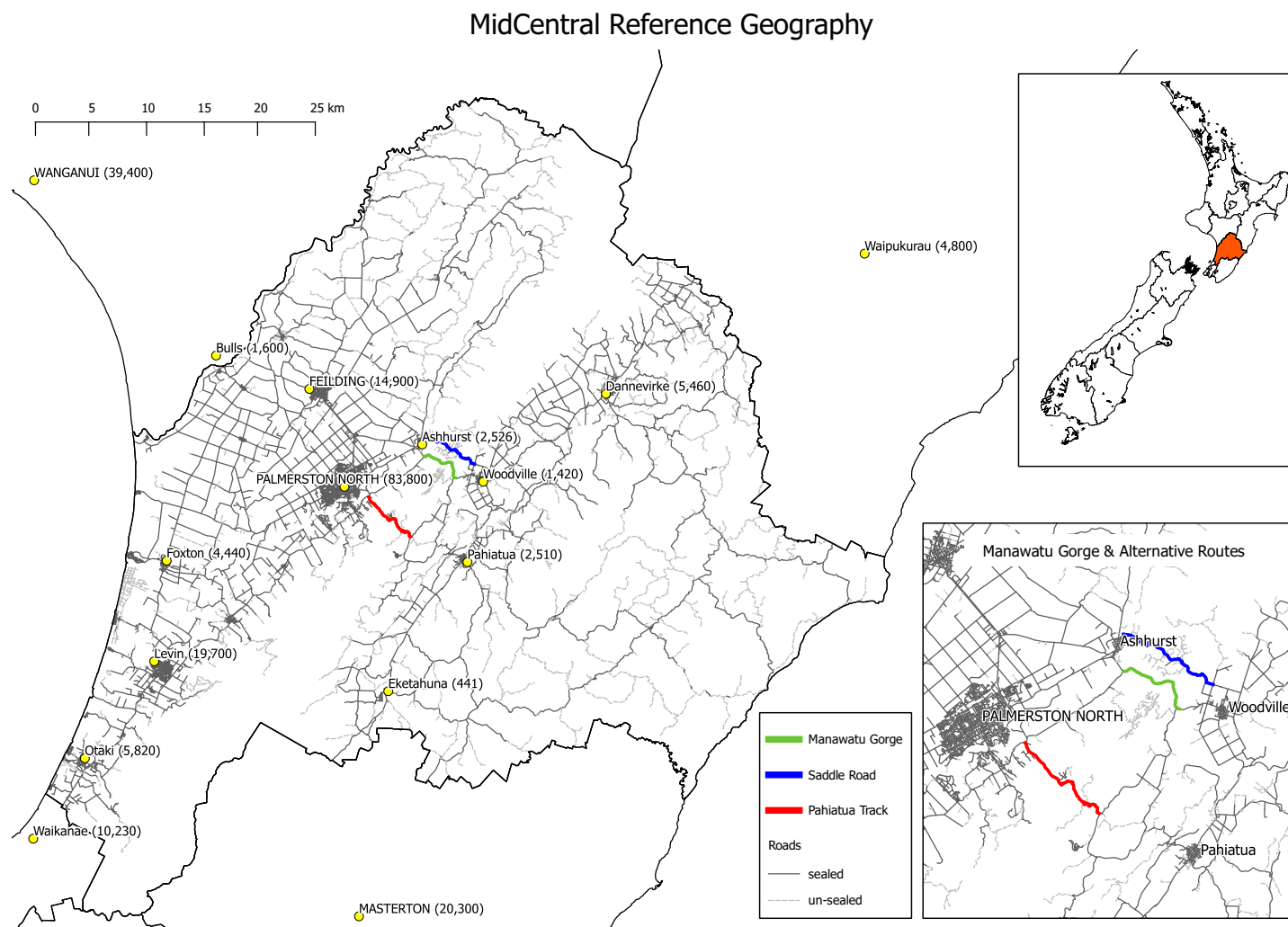


Figure 1.1: Reference Geography of the MidCentral District (urban area populations shown in parentheses)



Figure 1.2: *Aerial Photo of the Manawatu Gorge Slip (reproduced with permission of New Zealand Transport Agency)*

for it to be possible to detect service utilisation changes. A substantial population lives in the eastern area affected by the gorge closure. This population is composed of a combination of small urban settlements and surrounding rural catchments. It is possible to study how the same group of people used health services in circumstances with additional spatial barriers imposed; thereby reducing confounding factors caused by variation in population demographics and attitudes to health care.

The findings of this research may be useful specifically for the DHB when planning responses to future closures of the Manawatu Gorge. If negative impacts are seen in health utilisation during such a situation, regional health policy can be developed to address such impacts and reduce any associated disparities.

This thesis has a more general application for the planning and consolidation of health services. It will help to add to the literature by providing an insight into impacts on a homogenous population of individuals and their use of health services under different circumstances

of spatial accessibility. It may provide insight into the effects current health policy and the consolidation of services will have on utilisation in the future.

1.3 Research Objectives

This thesis aims to investigate a homogeneous population to determine if a difference in travel time affects their use of primary and secondary health services. It aims to complement the vast body of work that currently exists that highlights the importance of ensuring that health services are accessible to populations. It will help to inform policy makers on the potential impact current New Zealand health reforms may have on barriers to accessing health services.

The objectives of this thesis are:

- To provide a quantitative analysis of methods of spatial accessibility for measuring access to health services in a New Zealand context.
- To measure the change in health service utilisation for populations that have altered spatial accessibility due to the closure of the Manawatu Gorge.
- To investigate, develop and use novel geospatial information systems (GIS) methods to provide high fidelity data for analysis.
- To use a free and open source GIS software stack to compliment methodological reproducibility and to demonstrate open source's utility in such applications.

1.4 Ethics Approval

Ethics approval for this study was obtained from both the University of Canterbury Human Ethics Committee (approval number HEC 2012/152), and the Central Health and Disability Ethics Committee (approval number 12/CEN/21).

1.5 Thesis Structure

Chapter 2 begins by discussing the various factors that affect patients access to health services, including aspatial and spatial factors. It looks at the way in which spatial factors are calculated and methods and considerations for measuring both potential and realised access to services.

It then goes on in Chapter 3 to consider the way in which utilisation data can be used for measuring realised access in different settings.

In Chapter 4 I outline the methods that I used to improve the quality of the data with which access calculations were made. This includes the development and validation of both a corrective geocoding algorithm and a new travel-time model.

Chapter 5 outlines how the patient cohort was chosen for the study and how patients were allocated to different intervention and control groups. It also outlines the method I followed to calculate each potential and realised accessibility measure.

Results are presented in Chapter 6. These include data from the geocoding and travel time model validation as well as results of the analysis of accessibility.

In Chapters 7 and 8 I discuss the rationale, findings and implications of the methodological approaches that I have taken. I move on to compare the accessibility measures with each other as well as findings in other literature. Finally I consider the implications that these findings have for MidCentral and New Zealand. My final finding of the thesis is made in the conclusion.

Chapter II

Accessibility of Health Services

Health care accessibility is important for reducing health disparities in populations (Andrulis, 1998). Those that have the highest health needs are those that require health care the most. Poor health itself can act as an inhibitor to the ability to access health care and can also have a negative impact on a person's income. This can potentially feed into a negative-reinforcing cycle perpetuating health disparities between populations that are health and wealthy and those that are unwell and poor.

2.1 The Organisation of New Zealand Health Delivery

General Practice is the front-line medical care facility in New Zealand. General Practices are almost universally affiliated with Primary Health Organisations (PHOs). PHOs receive capitation based funding calculated on the number and composition of patients who enroll with them. General Practices receive this funding to help subsidise the consultation costs passed onto patients. The intention of this funding model is to make general practice responsible for maintaining the health of their enrolled population. Funding reaches general practice through District Health Boards (DHBs) and the Ministry of Health (MOH).

There are twenty-one DHBs in New Zealand. They are responsible for the health of the populations that reside in each of their defined geographical catchments. Each DHB has two divisions. Public hospital services are run by the provider division of each DHB while all public

health services both inside and outside the hospital are funded by the funding division.

Patients elect to be funded through a PHO by enrolling in an affiliated practice. There are no geographical limitations on where a patient lives (as long as they are a New Zealand resident) and where they can enroll. The cost of attending general practice is usually met by a combination of a fee paid by the patient, and a subsidy paid by the DHB. The cost to patients varies according to practices and patient demographics.

Patients do not pay fees for attending Public Hospitals in New Zealand. These are fully funded by each DHB and provide secondary health services. Tertiary services are provided out of hospitals in main centres. Hospital services are usually comprised of emergency departments, in-patient services, outpatient services and community based services.

There are some situations where patients live geographically distant from their nominated practice. This results in them being funded for general practices services by a DHB in which they do not reside. This situation occurs regularly in areas close to DHB boundaries, or in dense urban areas (such as greater Wellington).

There are a number of factors that influence the way in which patients use health services. Although public hospital services are free, and general practice services are highly subsidised there continue to be reasons that patients choose not to seek medical help when they may need it.

2.2 Factors that Affect Access

There are a number of factors that affect the way in which patients access health services. They can be categorised in a multitude of ways but for the purpose of this thesis I classify them into two groups; spatial and aspatial factors. Spatial factors are those that are related to the the geography of health services. Aspatial factors are those that are not directly related to geography. Factors may act by themselves to

either improve or hinder patient's access to services or they may work together in synergy or opposition to each other.

Factors that inhibit patients to attend health services are referred to as barriers.

2.2.1 Aspatial Factors

Financial Considerations

Accessing health services can cost patients money, both directly and indirectly. Direct costs are those that require patients to undertake a financial transaction that result in a net loss of money. Indirect costs are those where no financial transaction takes place but an opportunity has been lost to receive income.

Direct costs in New Zealand are limited mostly to primary care services where practices charge patients a fee to cover part of the cost of the consultation. In some areas these fees are waived for children under 6 or for those in the most socioeconomically deprived groups (Hefford et al., 2005). This is done to attempt to lessen the financial burden on attending general practice. Patients may incur costs for the dispensing and supply of prescription medicines, for laboratory tests and for radiological procedures.

There are no charges to patients for attending public hospital services. Diagnostic tests and medicines dispensed through hospital services do not carry the same fees incurred through primary care. They are funded through the total cost of the secondary care system and are not passed on directly in any way to patients. Private health care is available for some hospital services. The costs of these services are met by the patients and health insurers.

Other direct costs may be associated with with traveling to health services, such as public transport, petrol or parking costs.

Debt may discourage patients from seeking care (Beck, 1974). A patient or their family may have an outstanding bill for past services at a general practice. This in turn may make them less likely to seek care to avoid an embarrassing situation or to avoid the risk of confrontation over the debt.

There are indirect costs to seeking health care which are not so easily controlled. Most general practice services operate during normal business hours. Those who work who need to attend general practice, or take a dependent to general practice services must take time off work to do so. After hours services often charge a higher premium part-charge than general practices making such an option less attractive particularly to those in low income occupations. Anyone admitted to an inpatient hospital service will likely require time off work. Although there are provisions for sick-leave in New Zealand statues the need to arrange and take such leave may act as an inhibitory factor. All of these may play into the way in which patients seek health services (McIntyre et al., 2006).

Culture and Language

There may be cultural (Ellison-Loschmann and Pearce, 2006; Jansen et al., 2011) or language (Timmins, 2002) considerations that are factors in patients seeking care. Ethnic minorities experience less continuity of care in general practice (MacRae et al., 2006), which may be a result of the desire to find culturally appropriate care. There is evidence that Maori seek appointments in general practice more acutely and that they are offered less choice (Jansen et al., 2011). Having less choice may be associated with seeking care more acutely, where trying to make an appointment at the last minute may mean that there are fewer possible options to be given. Gaining an appointment at the desired time is the outcome of negotiation between patient and receptionist in most general practices (Gallagher et al., 2001). Cultural differences in negotiation strategies used by Maori patients may

be ultimately less successful.

Language barriers may discourage patients from seeking care. English is the predominant language used in New Zealand, and is the main language used within the health sector. Inability to understand language to a suitable level may make interactions more awkward when care is sought. Language barriers prevented 2.3% of those approached in a health survey from participating (Reeder and Trevena, 2003). This is a proxy of the level of language skills amongst the population where they may not be sufficient to have a basic conversation in English. It is likely that language barriers impact on patients health literacy.

There may also be some type of sub-cultural difference between those that live in rural and urban environments. Rural people are more stoic and have different health seeking behaviours (Jones et al., 1998). Where patients perceive their need for health care differently, some may wait longer for their conditions to resolve before seeking treatment. Some people may put up with more pain and discomfort, or loss of function than others.

Health Status

Having a long term illness or currently being in bad health both act as a barrier to accessing general practice and hospital services (Comber et al., 2011). The way in which health affects patients access to health services is varied. Patients may have difficulty with transport when they are unwell. There may be psychological factors where patients do not wish to seek care for fear of receiving bad news.

Patients with chronic conditions may find it more difficult to access health services because of factors associated with transport. Those that have their own transport may not be well enough to use it by themselves, and may rely on others take them to appointments. Those without their own transport may not be able to use public transport because of an inability walk to a bus stop or train service. Those with

a longer term illness may be off or out of work and may not be able to afford taxi services.

Patients suffering from chronic diseases may delay seeking medical help because they don't wish to risk hearing that their condition is worsening or that they haven't met their health goals (Webb et al., 2013).

Health Literacy

Low health literacy is associated with poor health outcomes (Baker et al., 1997; Berkman et al., 2011). The term health literacy refers to a patient's ability to comprehend and interpret health related information. While it may be related to language skills, the two are distinct. Where a person may be able to have a conversation and appear to have good language skills, their underlying comprehension of the language may not be sufficient to grasp subtleties and nuances often used in conversations. Health literacy is associated with patients' ability to read, their numeracy (McNaughton et al., 2013) and their ability to locate appropriate health information. Patients with low health literacy use health services related to preventative medicine such as screening and immunisation programmes less, while they tend to have higher associated hospitalisations and emergency care (Berkman et al., 2011).

Patients with poor health literacy may not associate early signs and symptoms of serious disease with the need to seek early care. Such aetiology is manifested in New Zealand's poor rate of rheumatic fever amongst Maori and Pacific Island populations; where sore throats in children are not seen as a serious health risk but when caused by the bacteria *Streptococcus-A* and left untreated can develop into rheumatic fever and may cause rheumatic heart disease later in life (Hale and Sharpe, 2011; Wilson, 2010; Jaine et al., 2008). Without sufficient health literacy there may be inadequate understanding of when it is important to seek care which represents a barrier to seeking appropriate care in a timely fashion.

Service Availability

The inability of patients to gain appointments in a timely manner can be a barrier to seeking primary care (Campbell, 1995). Primary care clinics have limited available space and limited health professional resource. Patients can be unwilling to wait for extended periods of time to be seen in a clinic and exhibit behaviours and strategies to improve their chances of getting appointments at short notice (Gallagher et al., 2001).

Patients who are not able to gain an appointment in their general practice may seek services in other general practices. They may attend after hours services or present at emergency departments. Aside from the already discussed financial implications for patients who present at another general practice or an after hours service there are issues with continuity of care. They are often not known at such practices and their full medical history may not be known. This can ultimately impact the quality of their care (Ekonomou et al., 2011) which in turn can impact their lifestyle or longer term morbidity.

Emergency Department (ED) presentations suffer the same continuity of care problems. Although presenting to ED has no direct cost to the patient, the cost to the health system is much higher than presentations to primary care.

2.2.2 Spatial Factors

Higgs (2004) identifies three categories of spatial barriers to health services; characteristics of the health system; the transport system; characteristics of individuals. The characteristics of the health system include service locations and the type and availability of resources. The characteristics of individuals include the individuals' locations and how they have access to transportation to move about their environment.

Health System Characteristics

The way in which the health system is configured spatially may affect peoples access to it. Services that are geographically diverse offer more people an opportunity to be close to one of those services.

Service location changes are a consequence of the current health reform policies in New Zealand. In the 1980s and 1990s New Zealand consolidated much of its secondary health system. This resulted in down scaling or shutting hospitals in smaller towns and centralising services to larger centres. The most recent reforms in primary care have Integrated Family Health Centres designed to be a consolidation of multidisciplinary services in one place (Cumming, 2011).

Not all facilities have equal resources. Resources for facilities may include consulting rooms, beds or staff. The ability of a facility to provide services is constrained by its most scarce resource. How resources are distributed among facilities and how close and accessible they are to populations can affect patients ability to use those services.

Transport System Characteristics

Patients who do not have access to private transport rely on public transport. Those who do not have good access to public transport may find it difficult to attend healths services that are more than walking distance from them. For those who are sick and infirm the distances at which services are accessible by walking may be very small.

New Zealand has among the highest car ownership rates in the world (Ministry of Transport, 2002; Auckland Regional Council, 2003). Approximately 90% of households have access to a motor vehicle (Statistics New Zealand, 2001). Motor vehicle ownership is a contributing component of the socioeconomic deprivation index calculation (Salmond et al., 2007).

The regularity and density of the public transport network will have an impact on how people use such services for access to healthcare (Lovett

et al., 2002). Because public transport in rural areas of New Zealand is poor (Pearce et al., 2006), those households that are rural and without access to personal vehicles may experience significant difficulties in accessing health services. Having no access to a motor vehicle has a negative association with general practice consulting for patients with asthma (Comber et al., 2011; Jones et al., 1998).

For those that live in urban areas, public transport is a much more viable option to assist in access to health services. Haynes et al. (2006) reported that 5% of patients used a bus to attend a hospital appointment post-diagnosis with cancer, rather than a private motor vehicle or taxi. This was the second most popular mechanism of attending.

If traveling to a health service by car, it is necessary to find an appropriate place to park one's car. The amount of time and difficulty that is anticipated in finding a car park may add to the perceived difficulty of a trip. Although the difficulty in finding a park is likely to be equal for patients who have short trips to health services, compared to those that have long trips, there is possibly a greater impact on those that live in rural areas that are not used to the rituals involved in finding an appropriate parking space in a large urban centre.

Haynes et al. (2006) found that the mean reported parking time for patients attending hospital services in the North of England was 7.4 minutes. This is difficult to factor into a travel time model because of the variation in parking times dependent on how busy a service is, available parking and tolerance for distance that can be parked from a facility.

Traffic conditions can be very different at various times of the day, days of the week, and weeks of the year. Particularly in urban centres traffic can be heavy in morning and afternoon peak times while people are transporting children to and from school and attending work. These may affect travel speeds by up to 25% (Messina et al., 2006).

Sinuosity is well documented in its use to calculate the degree to which lines contain bends (Heatwole and Lohani, 2004). It has been used

in applications to calculate how much rivers bend, as well as roads. Increasing sinuosity is associated with roads with more bends, or with bends which are tighter. The more sinuous a road the slower traffic will travel along it.

Any road condition that causes traffic to travel more slowly could be a factor which affects access to any health service requiring use of such a road. Traveling more slowly increases travel time. It also has a potential impact upon the perceived difficulty of driving the road.

Driving difficulty can affect the travel time on a road. A road that has more hazards, with more traffic, in bad weather, and is more sinuous will clearly take longer than a road without these impediments. There may however be a psychological factor in having to drive a road that is more difficult to drive. A relatively straight road that takes the same time to drive than a very sinuous, narrow road, with a great deal of oncoming traffic may be less mentally taxing to negotiate. This may affect drivers who are less confident at driving.

Travel distance to services has been an often cited barrier to patients accessing health services (Guagliardo, 2004). There are two main ways in which distance is measured; Euclidean distance which is a linear measure between the source and destination points; and Road network distance which measures the linear distances of each component of the road network that would be used to travel between the source and destination points.

Driving time is a product of the driving distance and the speed. People's perception of travel time is usually more natural and accurate than that of travel distance, where increased sinuosity of travel paths increases a perception of distance compared with that of time (Sadalla and Magel, 1980).

Patient Characteristics

The spatial area in which a person undertakes their daily activities, is referred to as their daily activity space (Nemet and Bailey, 2000). An office worker who lives within a block of their workplace, and has a supermarket and day care within another block, may have a very small daily activity space bounded by these destinations. A traveling sales person who shops in a different suburb to where he lives, and attends night classes may have a much larger daily activity space.

Daily Activity Spaces may be a factor in how people access health services in that it is a modifier of other factors (Nemet and Bailey, 2000). A person with a larger daily activity space, may tolerate (or appear to tolerate) a higher degree of spatial barriers. They may not consider distances or driving times the same as a person with a smaller daily activity space. Those with large activity spaces are presented greater opportunities to access services that are distance from their domiciles more than those with smaller spaces. This can have a particular impact on the elderly population (Cromley and Shannon, 1986).

How a person moves through their environment can affect how they access resources within that environment. Some patients with mental health issues have expressed reluctance to attend some clinics because of the need to travel through undesirable areas to reach such clinics (Cromley and Shannon, 1986).

For the very old, the young and those with physical impairment, mobility can be a large determinant of how easy it is to access health services (Cinnamon et al., 2008).

2.3 Spatial Accessibility

While the inverse care law makes statements about the proximity of health services to those that need them, it doesn't provide a method to calculate this objectively. Using spatial factors to account for populations of patients and service providers is one way to determine the

accessibility of health services. This concept is referred to as spatial accessibility of health services. There are however a myriad of ways which have been suggested for calculating spatial accessibility.

Access to health services may be expressed as potential or realised (Higgs, 2004; Gulliford, 2002). Specific literature use different terms to express each of these concepts. The basic definition and theme of each is consistent however. Potential access is that which could happen. In economic terms it expresses supply. It is the presence of enabling resources (Graves, 2009). Realised access is that which has happened. It represents service use or actual demand placed upon supply.

Calculation of potential access to health services has been the traditionally favoured approach. It is relatively easy to calculate because it requires only population level data on both patients and providers and because population level data is available through census and public data sets.

Realised access must be calculated using utilisation data. It is a function of the “Behaviour Model of Health Services Use”; people’s predisposition to use services, factors that enable or impede use and their need for care (Graves, 2009). The difficulty in calculating realised access exists in sourcing utilisation data and the additional computational resources that is required to make more granular calculations. Utilisation data must be received from provider facilities. As computerisation becomes more prevalent in services in the health sector, gaining access to such data sets is becoming more feasible. Utilisation data is in effect private health information, so there remain difficulties in obtaining these data sets because of privacy considerations. As computing power has followed Moore’s law and continues to double approximately every two years (Moore, 1965), the difficulty of making micro-calculations on utilisation data is decreasing.

The most basic is the patient to provider ratio.

Where resources are limited, the use of spatial accessibility measures can be used to model scenarios to understand the effect on popula-

tion subgroups of placing services is specific locations. In such situations, it is possible to target increased access to specific population sub-groups because these population sub-groups are not uniformly distributed across a geography (Christie and Fone, 2003). Understanding which sub-group needs to be targeted and which location from which to provide services to do that is one of the powerful uses of spatial accessibility measures.

There is a significant body of literature studying geospatial aspects of health care. The majority of this literature however investigates the potential for access to it, while very little looks at the realised access or utilisation of services in relation to spatial characteristics, and even fewer still look at outcome (Higgs, 2004).

2.4 Existing Literature

A number of studies use centroids to generalise patient populations (O'Reilly et al., 2001; Schuurman et al., 2006). This is likely due to difficulty in accessing patient level data. These centroids may be either geographic or population weighted. There are significant potential errors in using centroids, with a particular bias in rural areas. Because meshblocks vary in geographical size, and are designed to contain a roughly similar resident population, meshblocks tend to be larger in less densely populated areas. This situation is confounded by the fact that in rural areas the road network is also less dense, which can lead to even greater generalisation of calculations.

Lovett et al. (2002) is one of the few studies that has used patient level data. They did this using patient register information to provide a high detail construct of patient domiciles.

There is no clear definition of what is rural and urban land. Brabyn and Skelly (2002) used vegetation ground cover to determine those areas that were rural. Statistics New Zealand produces urban boundary information. This seems to encompass semi-urban areas.

Overseas census enumerations, similar to meshblocks have different characteristics, which can also make it difficult to generalise findings from overseas when aggregated measures are made. In Wales, 1991 census enumeration units have a mean area of 3.25 km² and in 2001 a mean population size of 462 (Christie and Fone, 2003). These differences are considerable in both size and population density.

There is also an assumption in most studies that people access services on an equitable basis. This is not true. The age profile compared with mean attendance rates, shows that older people are far more likely to visit their general practice multiple times in a year. People with more co-morbidities are more likely to use health services more. Those living in high deprivation areas (as defined at a meshblock level) are also more likely to have more health issues, and more co-morbidities.

2.4.1 Potential versus Realised Access

A significant proportion of research into geospatial measures of access to health care use measures of potential access (Higgs, 2004).

Jones et al. (1998) investigated utilisation of primary and secondary care services for those with asthma. There was a positive finding inversely associating travel distance to both services with service use.

2.4.2 The Nearest Service Fallacy

Some studies use the nearest service provider for calculations of travel time to services (Athas et al., 2000; Christie and Fone, 2003). Patients do not always attend their nearest service (Compass Health, 2012). There may be several factors that influence such a choice when patients consider health services. Service choices may be based on what are perceived to be suitable gender, ethnic or age profiles of service providers, or particular special interests that suit patients health needs. Choice may sometimes be motivated by desire to increase privacy by attend-

ing a practice other than that used by the majority of a community or family members (Haynes et al., 2003).

Travel time calculations are often made from patient’s domiciles to service provider locations. In larger urban areas, people may seek primary care close to their place of work rather than their domicile. Data sets which contain location information for patient’s places of work are more rare than those that contain their domicile.

Haynes et al. (2003) investigated the relationship between those registered with general practice in areas of England. They found that only 56% of approximate two million residents were registered with their nearest service provider. The frequency with which people attended general practice didn’t affect their choice, and neither did their ownership of a motor vehicle. Living in extremely close proximity to a surgery seems to positively affect patient’s choice to attend their nearest service however.

Hays et al. (1990) found that only 19% of residents in Gisborne, New Zealand attended their closest general practice. Proximity to services do however appear to play some part in choice, with patients in this study preferring services that they were near, just not necessarily the closest.

Assuming that patients attend their nearest service may be underestimating the travel burden placed upon patients.

2.4.3 Ecological Studies

Existing studies have tended to be ecological, relying on census or other population level data. Studies in New Zealand that have investigated spatial access to health care have used meshblocks. A meshblock is the smallest geographical unit for which statistical information is collected by statistics New Zealand, and in its current form was established in 1976, but has origins back to the 1891 census (Statistics New Zealand, 2013). Meshblocks are designed to contain approximately 70 to 110

individuals. This means that the areas of meshblocks vary inversely according to the population density.

Socio-economic deprivation calculations are based on information obtained about populations in each census although meshblock areas are re-calculated each year, . Often the last census meshblocks are used in health to ensure appropriate matches to the deprivation index. This can mean that calculations can be based on old or historic data.

The mean area of urban meshblocks in the MidCentral district is 15.6 km compared with 182.5 km for rural meshblocks. Their mean populations per meshblock are 95.7 and 46.8 respectively. The smallest urban meshblock is 0.009 km^2 with the smallest rural meshblock being just over ten times that of 0.1 km^2 .

In areas of less population density there are more generalised calculations being made.

Goodman et al. (1997) found a positive correlation between proximity to hospitals and hospitalisation rates, across all types of diagnostic groups, although they could not determine any causation.

2.4.4 Heterogeneous Populations

One of inherent problems of measuring the effect of travel time on people's health service seeking behaviour is that of having to study distinct populations of people. Many demographic factors that may influence health seeking behaviour such as deprivation and ethnicity can be controlled across distinct populations. When people are categorised based on their travel time to health service there will be a naturally higher proportion of people that live rurally in the more distant groups. Rurality and distance to health services are inextricably linked.

Rural people are more stoic (Fuller et al., 2000) and all factors being equal will tend to wait longer to seek help than their urban counterparts. Turnbull et al. (2008) measured the use of after hours telephone health service and found a decline with increasing distance from the

nearest primary care facility. If the barrier to accessing health services is distance and time, telephone consults should negate both of these barriers. Of course there may be other subtle psychological reasons that a person delays from seeking services via phone due to their geography. They may delay seeking help for fear of needing to present physically to a clinic, thereby re-introducing the geospatial barriers of time or distance.

This result does suggest however that it is extremely difficult to control for the effect of stoicism on health seeking behaviour because of the way in which it is specially correlated to distance to health services.

2.4.5 Geographic Generalisations

The most practical way to determine how long it takes a large number of patients to each travel to health services is to use a travel-time model that a computer can apply to such data sets. It would be impractical to collect actual driving times to various health services from a large patient sample. It would require a large amount of resource and it would be subject to inter-rater reliability issues. Some studies have used self-reported travel times to health services, one has shown little correlation between self-reported distance and actual distance (Siedner et al., 2013), another has reported high correlation between reported travel time and modeled travel time (Haynes et al., 2006).

There are two main methods of calculating travel time. Euclidean (also known as straight-line) travel times are based on calculating the straight-line distance and dividing by a speed factor. The second method is to use road networks to calculate travel distances more precisely and to use attributes of those networks such as speed limits, to determine the travel time for each segment of the network. Euclidean travel times are computationally easy to calculate and were therefore favoured more in early literature on travel times. Road network travel times have become more popular as computational power and access to geospatial information systems has become more ubiquitous.

2.4.6 Characteristics of Rural Versus Urban Meshblock Calculations

Table 2.1 shows a comparison of various metrics calculated to illustrate the characteristics that should be considered when trying to generalise calculations about individuals to meshblocks. The mean urban meshblock area is less than $\frac{1}{10}^{th}$ that of rural meshblocks.

Metric	Urban	Rural	Ratio
Meshblock Count	1,475	359	4.11
Total Usually Resident Population	141,201	16,803	8.40
Mean Usually Resident Population	95.7	46.8	2.04
Total Meshblock Area (km ²)	22,963.3	65,506.5	0.35
Mean Meshblock Area (km ²)	15.6	182.5	0.09
Smallest Meshblock Area (km ²)	0.009	0.1	0.09
Largest Meshblock Area (km ²)	1544.0	1560.4	0.99
Centroid to Furthest Perimeter Mean (km)	0.6	3.8	0.16
Road Network Length (km)	3,043.8	4,807.3	0.63
Road Network Density (m/km ²)	133	73	1.82
Road Network Intersections	4245	940	4.52
Intersection Density (intersections / km ²)	0.185	0.014	13.21

Table 2.1: Comparison of MidCentral DHB urban versus rural travel time calculation method metrics based on Census 2006 data

2.4.7 Euclidean Travel Times

A Euclidean distance is formed by measuring in a straight line between two points. It can either be measured linearly or computed using basic geometry principles from grid coordinates. It is therefore a good candidate to use from a computational perspective in large ecological studies with many permutations of source and destination nodes. Even the most modest computing platform today can calculate thousands of Euclidean distances per second.

It will almost always underestimate the actual travel distance and does not represent the actual path any individual would travel over anything but the smallest scales. Roads very seldom travel in straight

lines for any significant distance in New Zealand because of features like lakes, valleys, rivers and mountain ranges. Roads will tend to meander around significant geographical features. This effect has been identified in British Columbia, Canada (Cinnamon et al., 2008) and England (Jordan et al., 2004). Because Euclidean distances represent the shortest possible travel path between two points, using mean travel speeds alone would underestimate travel times.

Correlation coefficients between Euclidean distance and actual travel times provide the best way to calculate Euclidean travel times. Taking actual travel time measurements and correlating these to Euclidean distances will provide an indication of the strength of the relationship. If sufficiently strong, the correlation co-efficient can then be used to provide a travel time from any distance measurement. A correlation co-efficient will represent the relationship between the distance and both the discrepancy between the straight-line and Euclidean distance, and the mean travel time across the road network. It will vary as geography and road conditions vary. Geography that causes roads to be more sparse or be more sinuous would decrease the co-efficient. Road conditions that allow people to travel at higher speeds will increase the co-efficient.

There is mixed evidence about how well Euclidean distances correlate to actual travel times. Phibbs and Luft (1995) and Jordan et al. (2004) found good correlations between Euclidean distance and travel time under various conditions. Phibbs and Luft (1995) found a strong correlation for distances over 30 miles but there was a large variation in travel times for shorter trips with inter-quartile ranges of 25 minutes or 500% from the minimum travel time. By contrast, Jordan et al. (2004) found that sparse road networks and geographical barriers caused poor correlation between Euclidean distance and travel times for longer journeys. These mixed findings suggest that how well Euclidean distance can be used as a proxy for travel time may depend on local geography and road network conditions.

Euclidean travel times may be best used for collective travel times

across aggregated data (Phibbs and Luft, 1995). In a large enough sample the errors in individual calculations may cancel each other out. It would be appropriate to use such calculations in ecological studies. If applied to individuals however, Euclidean distance alone may be a poor estimator of travel time for each individual.

A small body of literature suggests that patient perception of drive time approximates linear distance rather than actual network drive time calculations (Haynes et al., 2006).

2.4.8 Road Network Travel Times

Road network travel times are usually calculated by determining a route by which a trip would take between two points, breaking the trip into smaller segments and using various attributes to calculate the estimated travel time across each segment, and summing the times across all segments. A prerequisite of calculating road network travel times is having road network information. The most basic information required is that which lays out the geospatial aspects of the road network, usually in two dimensions. Additional attributes may be available to help the travel speed calculations, including speed limits, road surface and lane information. It is also possible to derive additional attributes from the geospatial and existing attribute data.

Calculating road network travel times can be a computationally intensive task. The actual computations required depend on the exact methods and attributes used. Common steps in the process involve performing a least cost path analysis on the road network, to find the optimal route to use, then deriving attributes across the segments of the route, making travel-time calculations based on available attributes and summing the results.

Attributes that are used to calculate the speed with which a road may be traversed can be split into two categories; provided and derived. Many road networks are provided with attributes accompanying

the geospatial data. Road networks are commonly provided as a collection of two dimensional line geometries. Each line represents the road centre-line. Two dimensional line strings usually consist of two point coordinates. Each point represents the end of each line. Physical roads are often broken into numerous segments to allow straight lines to closely approximate road curves and to provide easy to interpret intersection information. Additional non-spatial attributes are often included. These attributes may include speed limits of roads, the type of material road surfaces are made of, such as tar-seal or gravel and special flags to indicate road classes such as motorways or state highways. These attributes are usually supplied for each line geometry in the network. Non-spatial attributes vary greatly between road network data.

Many attributes can be derived from road network data, or additional geospatial data. One of the most common attributes derived from spatial road network data is sinuosity. Sinuosity is a measure of how much a road bends. The technique for determining sinuosity is well established, being used for estimating characteristics of rivers and other geographical features where path curvature is of interest (Heatwole and Lohani, 2004). It is a relatively simple calculation. To determine sinuosity between two points, the actual length of the road segment is divided by the Euclidean distance between the same two points. Sinuosity calculations are sensitive to the resolution of the calculations being made (Heatwole and Lohani, 2004). As the road lengths used in the calculation increase, sinuosity decreases. This has an important implication when deriving sinuosity information in that road lengths should be kept constant in order to make comparisons between their respective sinuosities. Because the Euclidean distance is the shortest route between two points, sinuosity is always represented as a value greater than one. A value of one indicates a perfectly straight line with no bends. Beyond one, sinuosity is an arbitrary measure and individual values must be interpreted and used depending on need.

Another technique for deriving attributes for road networks is to use

additional geospatial data. An example of this would be calculating if a road was in an urban or rural area. Brabyn and Skelly (2002) used ground cover information to ascertain which roads lay in urban and rural areas. By deriving urban areas from ground cover information, it is then possible to perform geospatial calculations to find those roads that spatially intersect these boundaries and apply appropriate derived urban or rural road flags. An alternative to using ground coverage for this purpose could be using geospatial urban boundary information. Statistics New Zealand supplies such information as do other government agencies around the world for their respective countries.

2.5 Existing Road Network Models

2.5.1 New Zealand Models

The most comprehensive literature using a road network travel-time model in the New Zealand health system has been made by Brabyn and Skelly (2001) and subsequent related publications (Beere and Brabyn, 2006; Brabyn and Skelly, 2002; Brabyn, 2002; Brabyn and Barnett, 2004; Lauder et al., 2001). I refer to this model as the Brabyn-Skelly model (sometimes abbreviated to Brabyn).

Brabyn and Skelly (2001) used the NZ Topographical 50:1 (NZTopo50) map as the basis for their road network calculations. The NZTopo50 map has a limited number of supplied attributes. These include the road name, road surface, the number of lanes, a highway flag and one-way flag. Brabyn and Skelly derived a number of attributes including sinuosity, whether roads were urban or rural and whether roads were urban motorways. Sinuosity was calculated using road segment lengths of greater than 500m. They determined whether roads were urban or rural using a ground cover layer using a technique described by Thompson (1998). Motorway flags were added using road names that included the term 'motorway' and also by manual identification of roads in Auckland and Wellington urban areas with an open road speed limit.

Brabyn and Skelly (2001) classed roads into two categories; straight and bendy. They derived a sinuosity index for roads; those that had an index greater than 1.02 were put into the bendy class, while all others were considered straight. The choice of 1.02 as the threshold in this model appears relatively arbitrary:

”‘This threshold was determined by graphically viewing the sinuosity indices of roads in New Zealand and comparing this with the author’s personal experience.’”

From these derived attributes, eight classes of road were established. The way in which they were classified can be seen in Table 2.2. Each road class was assigned a travel speed ranging from 30 km/hr to 80 km/hr. The authors acknowledge a series of constraints that this model has, including the potential for incorrect assumptions where urban expressways allow travel speeds much greater than the 30 km/hr speed assigned to urban roads.

Sealed	Urban	Motorway	Lanes	Sinuosity	Speed (km/hr)
Y	Y	-	-	-	30
-	-	Y	-	-	80
Y	N	-	2	Straight	80
Y	N	-	2	Bendy	60
Y	N	-	1	Straight	80
Y	N	-	1	Bendy	40
N	-	-	-	Straight	50
N	-	-	-	Bendy	30

Table 2.2: Road classes defined in Brabyn and Skelly (2001).

Lauder et al. (2001) validated the Skelly-Brabyn model using New Zealand Automobile Association travel time information as well as empirical GPS data. They found that when compared with AA travel times, there was a high correlation between the model data and AA

data ($R^2=0.9616$) but that the Brabyn-Skelly model is approximately 6.6% slower. They reported that the GPS data was "‘much faster’" than the Brabyn-Skelly model, but do not report by how much.

Bagheri et al. (2005) use road-network travel time to calculate potential accessibility to health services in Otago. Although they do not explicitly state it, they appear to use the Brabyn-Skelly model, using derived sinuosity from road network data with arc-lengths of 500m, and classifying roads into eight categories with identical travel speeds to Brabyn and Skelly (2001). They don't state the sinuosity threshold they use. In a later paper the same authors use a similar model, but state they use a speed of 35 km/hr for the slowest road categories Bagheri et al. (2008). Neither paper provides any evidence of validation of the model.

Pearce et al. (2006) indicated that they used travel time across the road network to calculate indices of access. They stated that they used ARC/INFO network functionality which suggests that they used Least Coast Path Analysis. The model they used to calculate travel time was not stated however. Hays et al. (1990) in a study on a Gisborne population used road network distance rather than travel time.

2.5.2 International Models

Overseas literature is interesting to investigate only from methodological perspective. There is no evidence that overseas travel time models can be applied to New Zealand roads. New Zealand has its own unique combination of geography, speed limits and vehicle profiles and it appears that these factors may impact the performance of models (Haynes et al., 2006; Schuurman et al., 2006).

Christie and Fone (2003) used 48 classes of road with specific speed values for calculating travel time to hospital services in Wales. These included variables to account for four levels of population density being rural, urban, inner urban and conurbation central core. They also used a road type classification, a distinction between single and dual carriage ways and whether roads had passing lanes. This model used

speeds ranging from 19 km/hr through to 105 km/hr. They also used two adjustment factors of increasing and reducing speeds by 10 and 20% respectively for light and heavy traffic conditions. The authors state that they based their driving speed calculations on information supplied by a government agency, and that further local validation of the model would be needed for greater accuracy. This approach may be a simple alternative to much more complex calculations to model network demand and its impact on travel-time (Clark and Watling, 2005).

Cinnamon et al. (2008) used road atlas data which contained speed limit, road surface and stop sign information. The inclusion of stop sign information appears to be a relatively novel attribute available in the road network data in Canada. The authors state that their travel time predictions were accurate because of the attributes they used within the data set. They provided no evidence to support this position however.

Haynes et al. (2006) used a travel-time model with 16 classes of roads to validate travel times from patient domiciles to hospital services against patient self-reported travel times and Euclidean travel time calculations. Road classes were derived from an earlier study based on information from a government agency and adjusted from empirical test-drive data (Bateman et al., 1995). This study found that there was a high correlation between patient reported travel times and road network model travel times. There was no difference between Euclidean methods and road-network methods for estimating travel time. The authors acknowledge that the area studied did not have some significant geographical features that may affect Euclidean distance versus road-network calculations, such as mountain ranges or large water-ways.

Sauerzapf et al. (2008) used a travel time model as described by Haynes et al. (2006).

Schuurman et al. (2006) used the Road Atlas of BC data to calculate hospital catchments using a road-network model compared with a Eu-

clidean model. There was no clear pattern in catchment size. For some scenarios, the road-network provided a larger population catchment for services, while in others, the Euclidean model did. This result suggests that geography impacts travel time calculations dependent on the type of model being used and local variations in population structures.

Chapter III

Utilisation Metrics in Health Care

Utilisation of health services is a term used to refer to patients using health services. For general practice this means that a patient has had some form of encounter with that practice, which may be face-to-face, or via the telephone or perhaps even electronically through specially designed electronic messaging systems. For secondary services, this may be attendance at an the emergency department, outpatient clinic, or admission to a ward.

The term utilisation must be used carefully. Although it appears to be a simple and intuitive measure, the way at which utilisation figures are arrived at can be sometimes misleading (MacRae et al., 2006). Within the New Zealand setting Primary Health Organisation utilisation calculations exclude any encounters with general practice that are funded under any other stream other than directly by the patient or the PHO capitation funding. This excludes consults related to ACC, maternity and immunisations. Many PHOs run other special projects that fund services to patients which are also excluded from utilisation reports. PHO Utilisation data only include face-to-face contacts, unless the patient lives more than 20 kilometers from their general practice, when telephone consults can also be counted. For the purposes of this study, I intend utilisation to refer to all contacts with a health service that are face-to-face, including those funded by ACC and maternity services.

Studying utilisation in detail is more than just counting the number of attendances that may have been generated in a practice. Utilisation is multi-faceted. It may include an analysis of whether attendances are

planned or unplanned, how acute attendances are, and the types of interactions that occur. This may include the number of medications prescribed during encounters, or the number or nature of laboratory results.

Health Utilisation describes how health services are used. One of the most basic measures is the frequency with which a patient uses a health service. It is easy to observe and can be supplied in an intuitive measure for reporting purposes. Interpreting what utilisation changes mean however can be more difficult.

Utilisation rates can increase, remain the same, or decrease over time. What an associated increase or decrease may mean for a particular health service can depend on the context in which the change happens. When dealing with patients with chronic and deteriorating diseases, increases in utilisation rates are a natural expectation. As patients' disease states worsen they will tend to use health services more frequently (Gately et al., 2007). Public awareness of particular health initiatives may increase utilisation. Outbreaks of disease, such as the 2009 influenza pandemic can also increase utilisation. Changes in utilisation of one health service can affect the utilisation of others (Beney et al., 2009).

Because utilisation between health services is linked it can often make more sense to look at several different services as a whole. General practice services are considered to be a relatively inexpensive way of delivering health care and should be used to manage patients and keep them from entering hospital (Rosano et al., 2013). Emergency Department and Hospital services are expensive because they need to be staffed twenty-four hours as well as requiring a high staff to patient ratio. The commonly accepted desirable pattern between these services is to see general practice utilisation rise with subsequent falls in emergency department and hospital service use.

3.1 *Utilisation as a Metric*

Utilisation of health services is often used as an outcome measure for service effectiveness and access. Within primary care the goal is to have patients utilise health services to a high degree, as primary care is seen as much lower cost and preventative health delivery. By having patients attend primary care services frequently, the currently held belief is that patients will have their health managed more appropriately, and any problems will be identified early. The early identification of disease processes is important in preventing those processes from worsening in cases where they can be treated.

Rather than just focusing on frequency of utilisation, reviewing other aspects of it can provide insights for using it as an outcome measure. In a general practice that is proactive and has patients with well managed health, it is reasonable to expect to see a higher degree of planned presentations, and lower acuity.

The goal in secondary care is to have lower utilisation. Secondary services are usually expensive and their use is related to poorly managed health conditions. It is reasonable in Primary Care to expect that one wishes to see those people with serious health problems frequently, but those without less frequently. You do not want all of general practices' time being taken up seeing patients with low grade viral respiratory infections (colds), while being too busy to see poorly controlled diabetic patients.

3.2 *Emergency Department Attendance*

3.2.1 *Attendance*

Emergency Department care is expensive. People needing ED care are those in the most dire circumstances. ED presentations can be used as a proxy to well managed health status (e.g. asthmatics shouldn't turn up to ED if well managed).

Not all attendances to emergency departments require such a highly resourced and skills service to deal with their needs. In order to ensure the people with the most need are seen first in EDs in New Zealand, all patients are triaged and given a Triage Score. This score is a number ranging from one to five, with one being patients with immediately life threatening conditions and five being the least urgent and most minor cases. The triage score is intended to be used as a guide for how quickly emergency departments should treat patients in those triage categories. It can be used as a proxy for determining the appropriateness of presentation, although there is little consensus between health professionals in New Zealand about what is an appropriate ED presentation (Richardson et al., 2006).

Some people who present to ED could be seen in accident and medical (A&M) clinics or standard general practice. Accident and Medical clinics are general practice like services that operate longer business hours than a standard general practice. They often don't require appointments and have basic diagnostic medicine facilities on site. Those that are seen in an emergency department with a triage score of 5 are most likely to be able to be seen in the first instance in an A&M clinic or general practice.

There is a component of convenience in using an ED. It operates 24 hours a day 7 days a week, so can be accessed at any time in contrast to limited hours offered by most A&Ms and business hours offered by general practice. For those people who work, it can be a significant financial burden to simply take time off work to seek medical help. This can in a number of cases rule out seeking help at general practice. The cost of attending A&M clinics is usually higher than normal general practice fees. This fee structure can discourage people seeking treatment at A&M clinics, particularly when ED attendance is free.

Emergency departments can be seen as a one-stop-shop. They normally have access to diagnostic medicine, which provided through the hospital system is also free. This is in contrast to most A&M clinics which can normally have surcharge for X-Rays and other diagnostic procedures.

Emergency Departments have only one disadvantage to those attending them compared with general practice or A&M services. There can be significantly long waiting times for service especially for patients classed in a less urgent triage category.

3.2.2 Presentation Urgency and Triage Scores

Each patient who attends an emergency department is assessed on arrival and given a score from the Australasian Triage Scale (Considine et al., 2004). This score is used to determine the urgency with which the patient should be seen. The score ranges from 1 to 5. Those assessed as one are described as being “Immediately life-threatening”, while those with a five are “Less urgent, or dealing with administrative issues only” (Considine et al., 2004).

3.3 Hospital Admissions

The mainstay of hospital services are the inpatient services that are supplied in hospital wards. Patients are admitted to the hospital by several mechanisms. For patients who require extended care beyond what the Emergency Department can provide can be admitted to hospital directly from ED. Patients may alternatively be admitted to hospital for elective or scheduled procedures such as hip or knee replacement operations. Admissions end either through a patient being discharged from hospital, or they die.

The number of admissions that occur within a hospital are a function of demand and supply. The main supply capacity of inpatient hospital services is measured in beds. Beds are grouped into Wards. Wards are usually organised so that patients requiring similar medical or surgical specialties are admitted to the same ward. Available beds is an indicator of the number of patients at any one time the hospital can cope with in an inpatient setting. The ultimate ceiling of supply in hospital beds is determined by the physical space and actual number

of patients that can be physically accommodated. Hospitals usually operate at capacities below this upper limit if demand allows. This is often done to reduce the operating cost, as each open Ward has a minimum staffing level required, and staff are an expensive resource in hospitals. Operating below top capacity however does allow hospitals to adjust resources as demand requires.

Demand can at times exceed supply. This situation can come about when high demand and low supply peak at similar times. When this occurs it is possible for hospitals to alleviate this in-balance by delaying elective surgeries and therefore admissions. Where delaying elective surgeries is not sufficient to correct any in-balance, it is possible for hospitals to transfer patients to other facilities to free up local resources.

3.3.1 Length of Stay

The number of days that a patient stays in hospital is often referred to as 'length of stay' or LOS and is a common metric used for monitoring hospital performance. For the purposes of length of stay calculations, patients that die are not generally treated any differently in this metric from those that are discharged.

Some patients can be admitted and discharged on the same day. These admissions are usually referred to as day-stays, and consist of minor and relatively straight-forward medical or surgical procedures. There is no maximum length of time that a patient may stay in hospital, and this length of stay is generally determined by the severity of the patients disease, the amount of management they require, the complexity of their condition and how well they respond to treatment. Length of stay is a simple calculation to make, but because there are so many factors that contribute to the length of stay for patients it is little more than a rough proxy for understanding what is happening in terms of patient's and their case complexity.

Length of stay can be used as a proxy for resource use and to some degree case complexity. Patients who have long stays in hospital will gen-

erally be sicker or more complex than those who have short stays (Berki et al., 1984), and they will use more resources. Length of stay is however only a gross indicator and can vary greatly for similar conditions and case complexities particularly for older groups of patients (Martin and Smith, 1996).

3.3.2 Case Complexity

It is possible to use case-weighted discharges as a proxy for case complexity for hospital admissions in New Zealand. Case weights are calculated for every discharge from hospital for patients. The formula that is used in New Zealand is called the Weighted Inlier Equivalency Separation (WEIS). It is a rule based algorithm adapted from an Australian algorithm. It uses length of stay plus other variables to determine how many resources each patient used during their hospital stay. It includes considerations such as the patients diagnostic related group (DRG) and whether they had certain procedures that are particularly costly, such as mechanical ventilation.

Case Weights of higher value indicate a patient who have consumed more health care resource. This can be thought of as a proxy for case complexity. Patients who present with more severe disease will typically be more difficult to treat and will incur greater resources while in hospital, as a factor of their complexity and their length of stay.

Higher case complexity for those admitted to hospital can indicate a number of scenarios. The first scenario is that patients may delay seeking medical help in the form of general practice or attending an emergency department. This delay can lead to worsening of their disease state and subsequently greater difficulty in treating them. An alternative scenario is that patients may be well managed in general practice, with general practice able to deal with higher case complexities of patient disease states in the community. This in turn raises the threshold for those needing to be admitted to hospital.

Delaying seeking treatment in primary care can result in patients with

mild disease states that may otherwise be treated with relatively simple therapies, spiral out of control into severe disease states requiring hospitalisation. There may also be a lack of management of diseases by patients that can contribute to their conditions worsening.

If primary care manages patients well in the community, only the patients with the most severe disease states will get admitted to hospital. This will affect the case complexity in a way that those with minor issues that may otherwise have been treated in hospital are delayed being admitted and cared for in general practice.

In the scenario where case complexity demonstrates well functioning primary care, one would expect an associated decrease in the number of hospital admissions. The drop in admissions would be caused by greater numbers of patients being well managed in the community and not requiring hospital treatment. Where case complexity increases and hospital admission rates remain unchanged, or increase, this demonstrates an ineffective delivery of primary care to patients.

3.4 Outpatient Appointments

3.4.1 Attendance

Hospitals provide specialist services to the public through outpatient appointments. These are typically provided within the hospital, with patients having an appointment time and being seen by hospital staff in clinics. Outpatient appointments can result from referral to a clinic usually through a general practice or as a result of a follow up from a hospital admission.

Many patients will attend more than one outpatient appointment for any given episode of care. The outpatient appointments will typically be made well in advance of the appointment time.

Outpatient appointments can be important in the ongoing continuity of care for those with higher health needs that general practice can

provider. Follow-up care can aide in patients recovery, and may prevent subsequent exacerbation of chronic conditions.

3.4.2 Non-attendance

Some patients do not turn up to their outpatient appointment. These are colloquially referred to as DNAs, which is an abbreviation for Did Not Attend. There are two types of scenarios in which a patient may not attend an appointment; one in which they make contact with the service ahead of time and cancel or postpone their appointment; and one in which patients do not attend their appointment with no contact with the service. In some analysis these two types of non-attendance may be treated differently.

3.5 General Practice Utilisation

General Practice utilisation is made up of many different types of presentations. People may present when they are sick, for management of a health condition, for a repeat prescription or for screening purposes. Some people may be seen by a practice nurse only, while others may see a doctor.

Screening programmes are run through general practice. These programmes involve general practice keeping track of their funded population who are appropriate for the health screening. General practice will make contact with patients when they are due for screening procedures to be completed. These types of activities are not urgent. They may be delayed in times when practices are busy with acute presentations.

Those patients with long term conditions have regular contact with general practice to ensure that they remain supported and healthy. This happens even if they are in good health. During such visits patients will have lab tests ordered to aide general practice teams in understanding how well managed a patient's health is.

Utilisation in general practice represents a combination of acute and planned contacts. High utilisation rates by themselves cannot reveal how this type of presentation is split. Low utilisation however suggests that both acute and planned presentations are low.

Chapter IV

Method: Data Quality

4.1 Tools and Standards

Wherever possible the methodology of this study used Free and Open Source Software (FOSS). Data storage and geospatial calculations were done using PostgreSQL and PostGIS extensions. Network routing and least cost path analysis (LCPA) was undertaken using the pgRouting extension for PostGIS. Geospatial visualisations were completed with Quantum GIS. Geospatial data including road networks and place name information was used from the NZ Open GPS (NZOGPS) project and data sources made publicly available. All statistical analysis and graphing was completed using R and RStudio. Manuscript preparation was done using \LaTeX and TeXnicCentre.

PostgreSQL is a full featured relational database management system. It can store structured data in a large volumes and is efficient at matching and filtering data. PostGIS is software that extends the functionality of PostgreSQL by adding significant geospatial capabilities. These spatial capabilities can be applied to large data sets using the same efficient relational database architecture. PostgreSQL implements a procedural programming language called PL/pgSQL that allows users to create user defined functions.

Network routing functions are made available through the pgRouting extension to PostgreSQL and PostGIS. The pgRouting library uses the geospatial functionality and implements a number of algorithms used for least cost path analysis and graph traversal operations.

Quantum GIS is an open source geospatial platform that provides functionality to represent geospatial data graphically. It groups geospatial data into projects, with layers representing themes. These layers can be combined to produce visualisations from many data sources. Quantum GIS integrates with PostgreSQL leveraging the power of both systems.

R is an open source statistical programming language and environment. It is module in nature with many packages that extend its base functionality. Packages allow a range of calculations from simple to complex statistical procedures to be performed. It also provides packages to produce graphs. RStudio is an integrated development environment that wraps around the R environment and eases the integration and development of R programming scripts.

L^AT_EX is an open source type setting language and environment. It is designed for the production of scientific and technical documentation. It has several advantages over desktop word processing software that make it more suitable for the production of large technical documents. L^AT_EX itself is used to produce documents of varying formats, including portable document format (PDF) documents. TeXnicCentre is an integrated development environment that wraps around L^AT_EX making it easier to organise and provides features such as spell checking which are not inherently include in L^AT_EX.

Unless otherwise stated, all calculations of distance are made from the New Zealand Transverse Mercator 2000 (NZGD2000 / SRID = 2193). Longitude and Latitude data collected from geocoding and from global positioning system data are transformed using the PostGIS `ST_TRANSFORM` function from its World Geodetic System (WGS84 / SRID = 4326) representation.

Patient demographics were determined from the latest register information for each patient. Some demographics such as ethnicity can be recorded differently by practices over time. Apart from surveying patients there is no practical way of determining which demographic record contains the canonical information. Taking the last register

record is an assumption.

Age for patients is calculated as at the end of the study period.

Deprivation was calculated using the mean deprivation score from each patient's deprivation scores across the study period and converted back to a deprivation quintile based on the mean. Both gender and ethnicity were determined from the demographic details from 1 October 2012, representing the last submitted demographic details from each practice for funding purposes. Multiple ethnicity codes were ranked according to MOH guidelines, effectively giving preference to NZ Maori, then Pacific and then any other ethnicity over NZ European. Such ranking is done to ensure those with multiple ethnicities are not counted multiple times when statistics are grouped by ethnic code.

4.2 *Last-Best-Known Geocoding*

Using accurate patient domicile data is important in this thesis. The intent of this study is to investigate the effect that spatial barriers have on utilisation of health services. The chances of detecting small changes in utilisation increase with increased size of the study cohort. Patient's can only be included in the cohort if their domicile locations are known with reasonable accuracy because these are used to calculate the magnitude of spatial barriers, namely distance and travel time.

The source of patient domiciles in this study are the Primary Health Organisation register files. These are generated every three months synchronously on a national scale (MacRae et al., 2006). The only other nationwide data set of patients and their domiciles is the National Health Index (MacRae, 2006). This is predominantly populated from secondary care information. Most patients attend general practice at a far higher frequency than they do hospital services. PHO registers therefore provide one of the most up-to-date and frequent captured patient domicile data sets. Patient address data is captured as part of this data set. Precise address information can be converted into spatial

co-ordinates pinpointing patient domiciles.

PHO registers are geocoded by the Ministry of Health each time they are submitted. Domicile geocodes are calculated based on up to five address fields, although in practice, most patient management systems supply only three. The geocoding process results in the derivation of longitude, latitude and meshblock data where address data is sufficiently precise.

A meshblock is the smallest geo-statistical unit used in census level data and is most often used in health service analysis to associate individuals with a socio-economic deprivation. Meshblocks are designed to contain approximately equal numbers of individuals, usually from 70-110. Practically this is not always the case. Meshblocks can only be assigned to patients who have an address that is able to be geocoded with reasonable certainty.

Each address and geocode combination is given an uncertainty code as part of the geocoding process. The uncertainty code is used to indicate the accuracy of the calculation and ranges from zero to ten, with ten representing the highest uncertainty. Addresses with an uncertainty code higher than four are not matched to a meshblock. They are deemed to be too uncertain for practical use in quantitative analysis. Longitude and latitude information is supplied for addresses with uncertainty up to nine.

Uncertainty appears to change over time. Table 4.1 shows the results of this geocoding process for the same physical location expressed with slightly different address formats over different time periods. There is some temporal variation in the way in which the same address is coded and labeled as uncertain. In October 2009 the address that carries the suburb information is labeled with an uncertainty of 9, indicating a highly uncertain geocoding result, while the address without the suburb information is coded with a 2, indicating it is highly certain. In April 2010 both addresses are considered to be highly certain, with the address that includes suburb now showing a greater certainty.

Updates to the software that is used to perform this process may explain such observations. Updating street and suburb databases, or updating algorithms within the software could explain such changes over time. Continuous quality improvement initiatives in the geocoding process would be a reasonable explanation for this pattern.

Address	Date	Uncertainty
4/341 College Street, Palmerston North	October 2009	2
4/341 College Street, Westend, Palmerston North	October 2009	9
4/341 College Street, Palmerston North	April 2010	2
4/341 College Street, Westend, Palmerston North	April 2010	1

Table 4.1: *PHO register geocoding variation over time*

The MOH geocoding process does not always improve certainly over time. Table 4.2 shows the uncertainty and longitude and latitude values for the address 20B Denmark Close, Denmark Street, Dannevirke for five distinct quarters. Peak certainty is exhibited in January 2008 and degrades past that date. January 2008 has the best coordinate match compared with a manual check of this addresses location.

Date	Uncertainty	Latitude	Longitude
July 2006	9	-40.201278	176.106668
January 2008	0	-40.211514	176.095982
July 2008	9	-40.201278	176.106668
October 2009	10	-40.2106	176.099
January 2010	4	-40.210551	176.099149

Table 4.2: *Variation in PHO register geocoding certainty over time for the address 20B Denmark Close, Denmark Street, Dannevirke*

The most recent geocode from the MOH geocoding process may not be the most accurate. The implication of this for this study is the data

quality of patient register can be improved by applying data cleaning techniques.

4.2.1 Last-Best-Known Process

I have used a five step process to identify the most accurate geocode from the PHO registers. The intent of the process is for any given address to identify the most recent record with the best geocoding. The best geocoding is determined from the uncertainty codes where the most certain code is considered the best.

I loaded all addresses into a PostgreSQL relational database. Each address is represented by five fields of individual address components. For each unique address in all PHO registers I add a flag to each address record to indicate if it has an allocated meshblock. Addresses are checked for uniqueness and matched using all five address component fields independently (they are not concatenated together).

Those records having a meshblock are assigned a dummy-variable of 1 and those without 0. For each unique address, I find all duplicate addresses that have a meshblock assigned and discard those that don't.

From all addresses with a meshblock for each unique address I find the records with the minimum uncertainty value and discard the others. From the remaining addresses for each unique address I find the most recent record. I use this record to then update all other identical addresses with the associated geospatial, meshblock and uncertainty information.

4.3 Corrective Geocoding

Using the Last-Best-Known geocode can improve the data quality of the patient register for any address that is in a format precise enough for the geocoding process. Addresses that contain formatting errors or spelling errors may never geocode however. The Last-Best-Known

process described cannot generate a geocode for an address that has never been geocoded.

The geocoding process used by the MOH requires addresses to be separated into appropriate components that can be matched to known address information (Critchlow Limited, 2012). Often there are concatenations of address components into the same field, particularly in the case of rural addresses using the rural delivery 'RD' designator, or flat or apartment references. All of these variations appear to cause the geocoding process to fail. The fact that addresses must conform to a particular format and are subject to data entry errors has been long known to be a weakness of the geocoding process in PHOs.

General practice is provided with software that allows them to validate addresses when they enter them in their systems in real time. This is designed to minimise data entry errors that lead to poor geocoding by allowing operators to correct addresses immediately upon entry. Unfortunately these systems are not used widely due to perceived speed and usability issues.

I examined addresses in the PHO registers that were not geocoded by the MOH geocoding process. There appeared to be a finite number of reasons that addresses were not geocoded. Addresses were sometimes not actually addresses at all, at times they were not in the correct format, some had errors in them such as spelling errors and some had street or road names that didn't appear to be actual names. Some addresses had values such as 'no address given' or 'nil'.

I hypothesised that it may be possible to correct addresses with formatting or spelling errors sufficiently to match them to a road and address. Once matched it should then be possible to geocode them.

I have developed a process and associated information system to correct and geocode previously uncoded addresses. This process consists of five main tasks; converting road network data from a specific to generic format; loading of address data into a relational database for manipulation; standardisation of address components and information; matching

of address components to the known road network; and geocoding of matched addresses to specific coordinates.

4.3.1 Data Conversion

To geocode addresses it is necessary to have a database of road names, road types and associated geospatial coordinates. It is also desirable to have information on address number locations or ranges for the purpose of locating an accurate point on a matched road.

The NZOGPS project publishes road network and associated data which contains the attributes necessary for geocoding. This file is published in a 'polish' file format. It is composed of *sections* each having a number of *key* and *value* pairs. Each *section* encapsulates a concept or class of data, including cities, regions, and geometries. *Keys* represent attributes of each *section* with associated *values*. For geometry sections, these include road type, road name and geospatial co-ordinates. In some cases a *key* may have multiple data separated by commas, such as for street numbers and geospatial coordinates. The below is an example of a *section* representing a road segment.

```
[POLYLINE]
Type=0x6
Label=HADFIELD TCE
EndLevel=1
CityIdx=22
RoadID=929
RouteParam=2,0,0,0,0,0,0,0,0,0,0,0
Data0=(-41.29185,174.76257),
      (-41.29125,174.76263),
      (-41.29078,174.76278)
Nod1=0,14550,0
Nod2=2,453,0
Numbers1=0,E,22,4,O,21,1
[END]
```

This road segment has keys representing its name (Label), the city it is associated with (CityIdx), a unique identifier (RoadID), characteristics of the road and who can use it (RouteParam), road geometry data (Data0) and street number information (Numbers1). This format is intended for use in auto-routing devices such as in car GPS navigation. It is not initially suited for matching and geocoding purposes. It is however freely available, published under an open source license and updated several times a month.

The NZOGPS project provides data files for New Zealand in eight areas. MidCentral DHB is incorporated into the Wellington and Central files. In order to make it suitable for a geocoding task, I changed the format so each polish file was divided into four files encapsulating one class of data in each. These represent information for regions, cities, roads and street number ranges. The purpose of the conversion was to make the data more suitable for loading to and indexing in a relational database.

Format conversion was done with a custom-built software application, written in C#. This loads each polish file into memory, identifies each appropriate section, parses the appropriate attributes and writes the information to a comma separated value (CSV) file. City and Road data is linked together using unique identifiers. The software application must translate unique identifiers from each file into unique identifiers across all files loaded. This is because each polish file ID is only guaranteed to be unique for that specific file.

Each CSV file is cross-referenced to each other file using unique identifiers. Each row of the street number file has a reference to a corresponding road segment; each road segment to a corresponding city; and each city to a region. This is essential to ensure the integrity of the information when the separate files are loaded into the database. All files must match exactly for the data to be useful.

4.3.2 Loading and Formatting Address Data

There are three types of data loaded into the corrective geocoding database. In order to perform corrective geocoding the system needs a road network, geospatial information on NZ place names, and the address data to be coded.

The CSV files representing the road network are loaded into a spatially enabled PostgreSQL database. The database acts as the main storage for the road network, as well as the address data to be matched, and the corresponding match information once derived. PostgreSQL is useful because it supports efficient set matching operations of a relational database and provides geospatial functions through the PostGIS extension.

Each class of file is loaded into a separate database table. Cross-referencing information is maintained so that each database table relates to another by way of unique identifier values. Unique road information is derived from the road segment data based on the road name, road type, city and region attributes. The NZOGPS project uses some symbology to identify State Highways and this symbology is translated to its common textual representations (0xd~1 is translated to **State Highway 1** for example).

A file containing point coordinates and corresponding place names from around New Zealand is loaded into the database. This place name file is used during the matching process to cross-reference potential matches spatially with place names. It is sourced from Land Information New Zealand (LINZ) and is distributed under a creative commons license.

Address information for those addresses that cannot be coded are loaded from a CSV file previously exported from PHO patient registers. This file consists of five street name fields. Each field may contain a portion of an address. The addresses generally follow the convention of having street number and name information first, followed by suburb and city information.

4.3.3 Address Standardisation

Address standardisation is a process of taking addresses and ensuring that they conform to an expected format. This process involves removing redundant information, classifying relevant components and converting data to consistent representations. It uses computerised pattern matching and string comparison algorithms.

Addresses that are rural can contain a rural delivery number. This is used by the postal service to aide in mail delivery. It serves little purpose in geocoding addresses and is recommended to be stored as a separate address component. Often the RD number is included with another address component. When this happens the address will not geocode in the MOH process.

The algorithm identifies Rural delivery information using a regular expression. Regular expressions provide flexibility in pattern matching because they can match character classes rather than specific characters. Character classes are ways of expressing generalisations about characters. For example, instead of having to search for all types of characters that may represent a word boundary such as a space, full-stop, comma or a myriad of other characters, it is possible to use a single class to represent this concept. Using character classes is useful in address matching as patterns can be concisely built to account for digits and non-word characters that are often used in addresses, such as slashes, hyphens, or other punctuation. For example, the pattern below is what is used to identify and subsequently remove the rural delivery specifier.

```
(?<=\W)R\W*D\W*\d+
```

The regular expression is very short, but describes a complex search pattern. It can be described as 'find a match immediately following any space or punctuation character starting with the letter 'R' followed by zero or more spaces or punctuation characters, followed by the letter

'D', followed by zero or more non-word characters, followed by one or more digits'. Such a representation of a pattern concisely accounts for rural delivery numbers taking the form RD1; RD 1; R-D1; r.d 1; RD/1; or a myriad of many others.

Road type of each address is determined by comparison to a predefined list of road types. The exhaustive list of road types was generated from NZOGPS data. The match is done using direct comparison and also word distance comparisons. The last word of each address is checked for a case insensitive match to the list of known road types. If no match exists, it is checked for word distance to known road types using the DamerauLevenshtein distance algorithm.

The Damerau-Levenshtein algorithm calculates the minimum number of character substitutions, additions, deletions or transpositions that would need to be made to have the one word match another. This is referred to as a distance between the two words. Results with less distance represent words that are more similar to each other. This type of analysis can match words where one or both have a typographical or spelling error. This type of algorithm is used for modern spell-checkers in word-processing software being able to present options to correct spelling.

Words with a distance of 1 or less are assumed to match. I used the threshold of one because of the relatively small size of words used as street type designators. For example 'road' and 'lane' both have only four characters. As word distance approaches higher proportions of word length greater numbers of words match. At four letters and a distance of one, matching words share 75% in common. At four letters and a distance of two this drops to only 50%. A seven letter word with a distance of two shares 71% with its matches.

If no match is found a final step to guard against typographical errors omitting spaces between road names and the road type is done. To do this the algorithm checks the block of characters ending the address equal to the length of each known road type that is four or more char-

acters. If the distance between this block and known road types is only 1 or less then a match is found and assigned. A minimum size of four characters is used so that common road abbreviations such as 'RD' are not detected in genuine words ending in 'RD' such as 'FORD'. Road type descriptions are converted to consistent representations so that the many ways each type of road can be written are only done so in a single way for each type.

A road type designation is left unassigned for those roads where a match cannot be found.

Each address is then divided into its constituent components. These consist of road number, road name and road type. The algorithm identifies which of the five fields is the most likely candidate to contain the road numbers, name and type. Although most addresses are represented with this data first this is not always the case. The first field can contain descriptions of buildings or residences rather than actual address information. Sometimes apartment or flat numbers are stored in the first field by themselves. The algorithm identifies the most likely candidate by looking for combinations of numbers and street type descriptions.

A regular expression (see Appendix IV for implementation) is dynamically built based on the known road type, to divide the entire address into its appropriate components. This regular expression accounts for common field concatenations, such as the use of *unit*, *flat* and other non-standard address identifiers and their various expression permutations. A separate and slightly different regular expression is used for state highways. This is done because the format in which they appear is slightly different with a numeric suffix representing the road name.

The road name is encoded phonetically once it can be identified from the component analysis. Phonetic encoding takes a word and changes its representation into its constituent sounds. This is usually expressed using English alphabet characters. The result of this is that two words that are spelt differently but sound similar can be compared and found

to be equivalent. A simple example of this is the words 'here' and 'hair'. Both of these words would be represented as identical words after phonetic encoding. The purpose of this step is to eliminate issues created in interpreting words that sound similar, but also have value in negating spelling errors which can be caused commonly by substitution of characters that form similar sounds.

There are a number of well known algorithms for encoding data phonetically. I chose to use the double metaphone algorithm (Lawrence, 1990). It accounts for various foreign sounding names and words in common use in the English language, but remains relatively straightforward and is freely available.

4.3.4 Matching

Once addresses are parsed into their constituent components and standardised the algorithm attempts to resolve each un-coded address to a known location on the road network. To do this it works through seven methods of matching various attributes at different success thresholds. The methods are run in order of decreasing specificity. Once a match is found for each address the algorithm moves immediately to the next address. Table 4.3 shows the attributes and thresholds used for each method.

Matches are made between the address, road network and place name data on each attribute either exactly (E), phonetically (P), not at all (N), or within a range of values (R). The address attributes for road name, road type and road number are compared to the road network data using the matching technique for each method. Any road network data that has these attributes that match are a match candidate. Suburb and City information are taken from the address, matched to places in the place name geospatial layer. Euclidean distances are measured for each match candidate to matching place names. If the match candidate with the least distance between the match and place name is within the distance tolerance for that particular method it is flagged

Method	Road Name	Road Type	Road Number	Place Name	Tolerance (km)
1	E	E	R	E	5
2	E	E	N	E	5
3	E	N	R	E	5
4	E	E	R	E	30
5	P	E	R	E	30
6	P	E	R	P	30
7	P	E	N	P	30

Table 4.3: Matching methods with exact (E), phonetic (P), not matched (N) or ranged (R) comparisons

as a positive match.

4.3.5 Geocoding

Once addresses have been matched to roads the algorithm must calculate a spatial point to assign to addresses where a match exists. The spatial dimensions of roads are known but may span several kilometers. To estimate an appropriate point within these dimensions the algorithm uses address street numbers.

The algorithm uses one of three methods to geocode each address datum. Each road in the road network is composed of segments represented by straight lines. Each line has spatial coordinates indicating two points being each end of the line. Any given road can have many segments. The NZOGPS road network data contains a list of street numbers that span each road segment. The algorithm identifies the road segment within a road number matching the address road number. If it is not able to find an exact match, it finds a road segment with a street number range that encompasses a number closest to the actual address number. Failing these two methods it calculates the geometric centroid of the road and uses that as the matched point.

The algorithm must estimate where on a matched road segment a specific address is. The road network data contains only a list of street numbers that exist within each road segment. To determine an exact point along the segment where the address is the algorithm interpolates between the road segment start and end points. To interpolate a point the algorithm uses the formula below to calculate the proportion (A_p) to use. It divides the difference between the street number (A_n) and the segment address start number (A_s) by the address range. The address range is simply the difference between the segment end number (A_e) and start number.

$$A_p = \frac{A_n - A_s}{A_e - A_s}$$

The point at which the road segment is bisected in this proportion is the location used as the geocode for the address.

4.3.6 Validation

I validated the output from the corrective geocoding algorithm to establish how well it performed. I assessed the algorithm for its ability to correctly match roads and for those roads it matched its ability to specify an accurate spatial location. Gaining an accurate spatial location is dependent in the first instance on correctly matching the road.

Validation of the algorithm was done on a sample of records. The sample was stratified using the proportion of records matched by match method and those addresses that were not matched by the algorithm. The sample size corresponded to 1% of the overall un-geocoded addresses, 220 addresses in all. The minimum size of each strata was 10 addresses. None of the records used in the validation sample had been used to develop the algorithm to avoid training bias.

Each address in the validation set was manually assessed. All address data was reviewed and the road network was manually searched for each address to find a match. The corresponding road network identifier

from the road network was recorded against the validation data set. Some addresses were not possible to geocode, as they contained non-address entries, such as 'needs new address' or contained only PO Box numbers. These addresses were marked with an invalid flag. This process took approximately 3 hours.

Each address that was manually matched was then geocoded using the Google geocoding API. The API was passed the correct road network name based on the road network road attributes determined through the manually linked road network identifier. The API returned a latitude and longitude. I used Google Maps to manually geocode any address that did not geocode via the API.

I calculated the Euclidean distance between the corrective geocode algorithm and that determined from Google. I considered this to the error of the algorithm.

4.4 Developing a Travel Time Model

I set out to use the New Zealand GPS Open data to establish if I could improve upon the performance of the Brabyn-Skelly road network travel time model (Brabyn and Skelly, 2001).

4.4.1 Existing Attributes

The NZGPSO data extends the attributes available from the New Zealand Topographical 50:1 (NZTopo50) data set seen in Table 4.4.

The speed limit attribute provides speed information in the range of 20 - 110 km/hr. The intention of this attribute is not to indicate the legal speed limit, but rather to indicate a proxy for the estimated travel speed across that particular section of the road network. For example, an urban street that has a legal speed limit of 50 km/hr may not be able to be traveled across at this speed because it has speed bumps. The intention of this field is to provide GPS navigation units with

Attribute	Degrees of Freedom	Minimum	Maximum
Speed Limit	7	20	110
Road Class	4	1	4
Road Type	6	1	6
Toll Road	2	0	1
Not For Car	2	0	1

Table 4.4: Additional attributes supplied with NZ Open GPS data.

some proxy of travel time for various routes so that it may estimate the fastest route.

The road class attribute is included to allow older GPS units to use a smaller range of calculations when making routing decisions. When combined with the speed limit class, it may provide an additional level of granularity to the speed limit data.

The road type attribute provides information on the physical type of road. This includes the road surface and number of lanes available. There are three main types of road surface on New Zealand roads; sealed, metaled and non-metaled. Sealed roads are those that you expect to see on main and urban road networks and are composed of a hard surface, usually consisting of chip-seal or bitumen. Metaled roads are those that contain shingle or gravel as a surface and non-metaled are those that are usually grass or dirty only.

The NZOGPS data also contains a number of Boolean flags. In this case, both the "toll road" and "not for car" attributes have values of 0 (false) and 1 (true). Some roads are not for private motor vehicles (service access lanes as an example), and these have the "not for car" flag set to 1. This allows routing algorithms to ignore these roads in LCPA.

4.4.2 *Deriving Journey Segments*

Any journey on a road network from a start to an end point can incorporate roads segments with varying attributes. Each road segment is of varying size. Some attributes derived from the spatial properties of the road network are dependent on geometries to be of similar sizes so that they may be applied to travel-time calculations consistently.

To accomplish this for any journey I concatenated each contiguous road segment to form one geometry. From the journey start point I divided each geometry into journey segments of exactly 500m. The final journey segment of each journey was a variable length. Existing attributes of each journey segment were applied using a weighted mean based on the existing attributes of the constituent road segments. Attributes which are categorical were assigned by giving each category a numerical representation, performing a weighted mean operation on those numeric values and converting the mean back to the category which it most closely represents. Calculations were completed in PostGIS with a user-defined function.

Derived attributes were calculated across each journey segment.

4.4.3 *Deriving A Sinuosity Index*

A sinuosity index (S_i) may be calculated for any line geometry by dividing the length of the geometry (l) by the Euclidean distance between the start and end points (d_E).

$$S_i = \frac{l}{d_E}$$

Indexes range from 1 representing a perfectly straight road where $l = d_E$. They cannot fall below 1 because a Euclidean distance represents the minimum distance between two points. Typically sinuosity indices are calculated in two dimensions only.

The sinuosity index is sensitive to the length of line geometry it is applied to. For geometries of greater length the index is shifted towards 1. Journey segments except the final segment of any journey were of consistent length.

4.4.4 *Deriving Weighted Mean Sinuosity Index*

A weighted sinuosity index (S_{wi}) was calculated for each journey segment. This consisted of using the sinuosity index of the journey segment before (S_{ib}) and after (S_{ia}) the current segment and combining it with the sinuosity index of the current segment (S_i).

$$S_{wi} = (0.25 \times S_{ib}) + (0.25 \times S_{ia}) + (0.5 \times S_i)$$

The current segment contributes half of the weighting and the before and after indices contribute 25% each. For segments at the start and end of the journey, the weighting was applied with the current segment contributing 75% of the weight.

4.4.5 *Deriving Intersection Density*

Network intersections were determined by counting the number of road segment vertices in the road network. Each road network segment has two vertices representing its start and end points. A vertex is represented once for each road segment it is a part of. The frequency of vertices in a road network describes the characteristics of the road segments with respect to how they join other road segments.

A vertex that occurs only once indicates the end of a road segment with no continuation, commonly referred to as a 'dead end', 'no exit' or 'cul-de-sac'.

A vertex that occurs only twice indicates a road segment that is continuous with one other road segment. In this scenario vertex appears in the network twice, once to represent the end of one segment, and

once to represent the beginning of another. Vertices with a frequency of two are common because of the way road network is represented by a large number of small road segments.

A vertex that occurs more than twice in the road network must represent an intersection. The number of roads forming the intersection is represented by the frequency of the vertex.

I used the road network segment of vertices to identify the spatial locations of intersections. I placed a 500x500m grid over the road network. I chose this grid size to be consistent with the journey segment length used for derived attribute calculations. I then counted each road intersection that was present in each grid. I examined visually the pattern of grids with 2-5 intersections to determine how they corresponded to urban boundaries. Those grids with two or more intersections I classed as high intersection density.

Where a vertex lies on the boundary of more than one grid, it was counted once in each grid.

4.4.6 Deriving Elevation Data

The NZOGPS project data does not contain elevation information for the road network. Neither does the NZTopo50 road network data. In order to include elevation data into the travel time model I first needed to find a source of data that would contain such information.

The only elevation data I was able to source for the Manawatu Gorge area was twenty meter contour data. This data contains elevation iso-bars for each twenty meter step in elevation above sea-level. Because twenty meter elevation changes are relatively large, these contours are not correlated well spatially with the road network. Although higher resolution data is available in contours of 1 meter or below using LiDAR measurements, those that are available tended to be for urban areas, or areas where there is a need due to emergency or planning scenarios (for modeling tsunami impact as an example). Because the

Manawatu Gorge and surrounding areas is rural such high resolution data does not appear to be not freely available.

I calculated road elevation using a simple algorithm that used the elevation of the closest contour to each road segment centroid. Using PostGIS I calculated the centroid of each road network segment, and then used a nearest neighbour function to identify the contour line closest to this centroid. In the event of two contour lines being equidistant I arbitrarily took the first contour record by ID.

Additional to this method, I also designed an interpolation algorithm to attempt to provide a more graduate elevation profile. This operated on the same 20m contour data and used road segment centroids, nearest neighbour calculations and next-nearest-neighbour calculations to provide a weighted mean elevation. A full description of the interpolation methodology can be found in the Appendix A.

Validation

To evaluate the interpolation model, I used a Garmin Nuvi GPS unit to record the x,y and z coordinates of the road network for the Manawatu Gorge. Two-hundred and twenty eight observations were made over the 10.8 Km journey, equating to an observation every 47 m of the trip. The logged trip data was transferred from the Garmin GPS unit and transferred to a PostGIS database for analysis.

4.4.7 Deriving Maximum and Total Turning Angles

I calculated the angles for all turns in a journey segment. I did this using a user-defined function in PostGIS. Each turn in a journey segment is represented by three vertices. The function receives three point geometries, one for each vertex in the turn. The first point represents the origin of travel, the second point represents the point at which direction of travel along the journey segment changes, and the third point repre-

sents the destination of travel. It calculates the angle formed between points and returns this as the turning angle.

Each journey segment is made up of two or more vertices. The number of turns in a journey segment can be represented by subtracting two from the number of vertices. A journey segment with two vertices represents a perfectly straight line.

For each journey segment, turning angle calculations are made for each turn. The turning angles for each journey segment are summed and stored as the total turning angle attribute. The single highest turning angle is stored as the maximum turning angle.

4.4.8 Modeling

Empirical Data Collection

I collected GPS data using a Garmin Nuvi by driving a variety of roads in the lower North Island of New Zealand, bounded in the north by Napier and Bulls and in the south to Wellington. These roads included a number in the MidCentral district and the Saddle Road and Pahiatua Track. Data collection was done on nine different days of the year. This GPS data comprises 12,374 individual observations over 557 kilometers of unique roads of differing classes. Each observation contained x,y and z coordinates,date and time and a sequencing identifier. Each GPS observation represents a geospatial point.

Each point of each trip was joined in sequence to form line string geometries of not less than 500m. Existing road network attributes were assigned to the GPS line strings by using nearest neighbour calculations to road centre-line data. Derived attributes were calculated for each journey segment. The observed speed for each journey segment was calculated using GPS time stamps associated with each observation.

Multiple Linear Regression Modeling

I used the observed GPS data including existing and derived attributes in an Akaike Information Criterion (AIC) stepwise regression to identify the most promising attributes to contribute to a linear model. The variables used in the regression were speed, road class, sinuosity index, lanes, log transformation of intersection density, high intersection density flag, total turning angle, maximum turning angle and an urban flag.

Road class and lanes were transformed into dummy variables representing each category of each attribute. The regression algorithm used a backward search pattern with a maximum number of steps of 1000.

Speed Constraint Rule Model

I developed rules to calculate speeds using attributes to determine speed constraints. Both upper and lower constraints were set based on journey segment attributes. I examined each attribute by magnitude and observed speed by each speed type. From patterns in these observations I developed rules for attributes which defined maximum and minimum speed constraints.

I developed a set of rules that were intended to determine the traveling speed across a journey segment by using attributes to determine which provided the maximal constraint upon speed. Attributes that used continuous variables were converted to ranges and used as classes. From these constraints, a speed for the journey segment is determined.

The full implementation can be located at the address listed in the Appendix B-1. Each rule begins by constraining the speed using the speed information to set a lower and upper bound for speed. Rules that constrain the upper and lower bounds are then applied in order of least constraining to most constraining.

```

if speed_class = 4 then
  lower := 80;
  if lanes = 4 then lower := 80; end if;
  if lanes = 3 then lower := 75; end if;

  upper := 100;
  if sinuosity_index > 1.1 then upper := 95; end if;
  if road_class = 0 then upper := 90; end if;
  if max_turning_angle > 200 then upper := 75; end if;
  if sinuosity_index > 1.4 then upper := 80; end if;
  if road_class = 1 then upper := 60; end if;
  if sinuosity_index > 1.8 then upper := 55; end if;

  factor := 0.9;
end if;

```

The way in which constraints are applied differs for each class of road speed limit. Not all attributes are used for all speed classes, only where there appeared to be some constraining effect in either an upper or lower direction. The rules were derived from looking at the correlation between speed and each attribute at each speed category.

An adjustment factor is set for each speed class. The purpose of this adjustment was to bring observation data back to a likely median driving speed of the population. The observation data suggests that the speed at which observations were made at times exceeded the legal driving limit. The factor was set to bring maximum driving speeds for each road class down below the legal driving speed limit.

A final speed for the journey segment is determined by taking the mean value between the upper and lower constraints and applying the adjustment factor. For ease of reference I will refer to this model as the MacRae model.

4.4.9 Model Validation

Model Comparisons

I compared four travel time models; the Brabyn-Skelly travel time model (Brabyn and Skelly, 2001); a Basic travel time model; the Google Maps model; and the MacRae model as described in section 4.4.8.

The Basic travel time method was used as a reference standard in this analysis. It involves using a calculation based purely on NZOGPS speed type data. Speed type data is contained within the NZOGPS data which is used to form the road network for these analysis. The Basic model is not intended to represent actual travel time, but is used as a reference and intuitive measure of travel time.

I implemented the Brabyn-Skelly model as a PostGIS function (Appendix B-2).

The Google model represents times calculated from the Google API for the primary suggested route. Google Maps is a free service provided to the public. Details of how Google calculate their travel time information are difficult to find, but it is a widely used and understood service. It is included for the purposes of comparison due to people's familiarity with it and the results it provides.

In order to compare model performance, 30 locations were chosen from MidCentral that represent a range of travel times from short trips of only a few minutes to long trips beyond 60 minutes. The trips were chosen to represent both eastern and western areas of the district. All trips used the same destination point of Palmerston North hospital.

Each source and destination trip was computed in each model. The method used to calculate times using the Brabyn-Skelly and MacRae model is described in section 4.5.

Traffic Survey Data

In order to validate and compare models against real-world performance I conducted traffic surveys to measure the travel speeds of the public across selected roads. The two alternative routes to the Manawatu Gorge were included although at the time the data was being collected, the Manawatu Gorge was not open so was unavailable for surveying. The Rimutaka Hill road was included because of its special mention in by Brabyn and Skelly (2002) in considering attributes for their model.

Routes were chosen to optimise the number of observations available within short windows of time, as well as to minimise the number of vehicles that would stop or take alternative routes between the start and end points of the measurement. These conditions meant that no unsealed roads were surveyed due to the low traffic volumes that would be encountered on such roads. The criteria to minimise routing options for vehicles also meant that the observations were limited to short distances and precluded any dense urban road networks. A mixture of straight and sinuous roads were included.

A video camera was positioned at each end of the chosen road segments and traffic entering and exiting the segments were recorded. Each video was time-synchronised using a visual or audio marker. The videos were reviewed in tandem at a later date and the exact entry and exit times of each vehicle were noted as well as their direction of travel.

Data was collected during differing weather conditions and different times of day. Data was not collected for hours after dusk or before dawn due to the practical limitations of being able to accurately identify vehicles in each video sequence. A total of 450 observations were made across 6 sites.

4.5 Calculating Travel Times

4.5.1 Road and Routing Network

All travel time calculations were made on the NZOGPS road network data. This was converted from polish file format and loaded into a PostgreSQL database using the same technique described in section 4.3.2 (page 55). Each road in the network was converted into line geometries. Each line geometry was converted to two point geometries representing each end of the line. These point geometries were matched and converted into a routed network schema.

I combined the attributes from the NZTopo50 network with the NZOGPS data by matching the road segments in each network together using a nearest neighbour calculation. New Zealand Open GPS road segments were assigned the attributes from the NZ Topo50 data that was closest to them.

4.5.2 Least Cost Path Analysis

Least cost path analysis takes a network and associated costs and calculates the path with the least cost to traverse the network from a source to destination point. The Dijkstra algorithm has been a fundamental basis for LCPA calculations since published in 1959 (Dijkstra, 1959). The algorithm stores the costs for each combination of path in the network until the destination point is reached. This has the result that once a path reaches the destination there are no other combination of paths that can reach the destination with a lower cost. The Dijkstra algorithm performs well in small networks or for small volumes of calculations but because of the geometric expansion of the algorithm to find all path permutations, performs less well on larger networks such as real road data.

The A* (pronounced A-star) algorithm is an extension of Dijkstra. It uses the same basic principle but uses heuristics to grow the most

promising path permutations before the least promising ones. Because of this it is able to locate least cost paths with greater efficiency and speed than Dijkstra. This has a particular advantage in large networks, such as road networks. There are numerous other least cost path algorithms, each with increasing complexity and computational requirements. Dijkstra and A* are implemented consistently across many GISs.

I used the A* algorithm because of its performance metrics and because it was available in PostGIS as part of the pgRouting library.

One limitation of the A* algorithm is that the cost of each routing segment must be used as an input to the algorithm and therefore already known. The Brabyn-Skelly and MacRae travel time models both determine travel times based on attributes derived from the geometry of paths (e.g. sinuosity index and turning angles). This represents a problem because such attributes cannot be calculated until a path is known, but the least cost path cannot be calculated without cost data.

To solve this problem I used a two-step LCPA procedure (Figure 4.1). Square boxes indicate process related activities, while curved boxes represent stored data passed from one process step to the next.

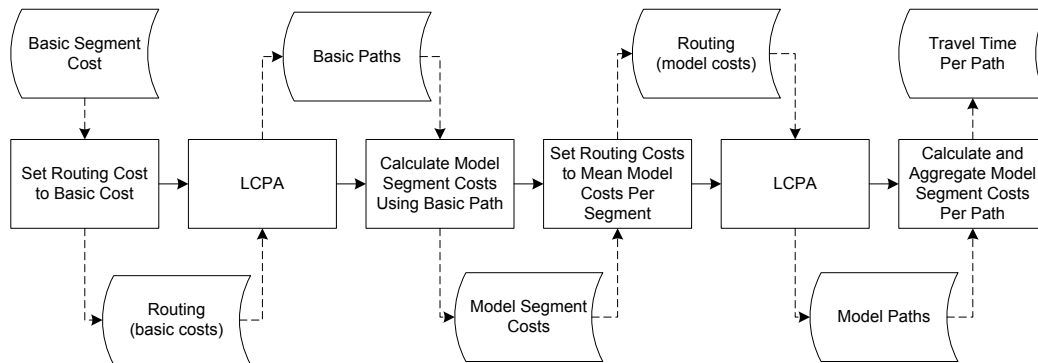


Figure 4.1: Two-Step LCPA Procedure

I first set the cost of each routing segment to that derived from a Basic travel time calculation. The basic cost is a simple calculation that can be made on the basis of existing attributes associated with each

road segment. The basic cost (C^b) is the travel time derived from the segment length (l) and speed limit (s):

$$C^b = \frac{l}{s}$$

I then undertook a LCPA on the routing data using the basic costs. This resulted in a path for each source and destination node pair. Node pairs consisted of patient domiciles and health service locations. I then used each path in the travel time model and stored the resultant model segments and speeds. I calculated a direct model cost (C^{dm}) for each routing segment by using its length (l) and the mean speed from spatially intersecting model segments (\bar{s}^m) :

$$C^{dm} = \frac{l}{\bar{s}^m}$$

I then set the cost for each routing segment to the newly calculated model cost.

Not all routing segments spatially intersected with a model segment. Those that didn't couldn't have a direct model cost calculated due to the absence of model speed data. I set such routing segment costs to an indirect model cost (C^{im}) by adjusting the basic cost by the mean ratio between all paired basic and direct model segment costs where a direct model cost has been derived:

$$C^{im} = C^b \cdot \frac{\sum_{i=1}^n \left(\frac{C_i^{dm}}{C_i^b} \right)}{n}$$

I then completed the second LCPA with the new routing segment costs and stored each resultant path. These paths then represented the least cost paths based on an estimate of the model travel times using derived attributes. Finally I calculated the travel time for each path.

The purpose of the two step LCPA process was to provide a more accurate representation of the routes people are likely to take between two

points to minimise traveling time. Using the traditional one step LCPA resulted in finding the path of least cost based on existing road segment attributes. A more sophisticated model using derived attributes can then be applied to those paths to determine their accurate travel times. The calculated travel times are predicated on the assumption that the path used was indeed the fastest. This assumption will often be wrong due to the LCPA process only using existing road segment attributes and not the complex and descriptive derived attributes. The two step LCPA process uses more accurate cost information derived from the desired model.

4.5.3 Gorge Closure Simulation

The purpose of this study is to understand the impact the closure of the Manawatu Gorge had on utilisation of health services. The initial LCPA was done on a road network that had the Manawatu Gorge open. It was necessary to calculate the travel times for people once the gorge was closed.

The routing network was modified to remove a single road segment from within the gorge. This breaks the continuity of the road and prevents any LCPA routine from being able to complete a trip through the gorge. Because the Manawatu Gorge road has not intersections or alternative paths along it's course it was not necessary to remove the identical segment affect by the slip that resulted in the gorge closure.

People using the Manawatu Gorge were identified by identifying from the path information generated from the initial LCPA any person that had a road segment that was a part of the Manawatu Gorge road in their route. Trips to the hospital or their GP were flagged respectively.

The same two-step LCPA procedure was performed for the subset of people flagged as originally use the Manawatu Gorge. The LCPA was conducted using the modified routing network. The travel time was stored as an alternative travel time value for hospital and general practice.

4.5.4 Final Calculations

Each patient had travel times for the Brabyn-Skelly, MacRae and Basic travel time models calculated with the Manawatu Gorge open and closed from their domicile to PN Hospital and their funded general practice. Patients using the gorge were flagged by identifying those who had a travel path that contained a road segment identified in the Manawatu Gorge.

Chapter V

Method: Patient Cohort and Accessibility Measures

5.1 Identifying the Patient Cohort

The way in which the cohort of patients was included in the study is shown in Figure 5.1

Patient geocodes were updated with their last-best-known code (see section 4.2 on page 48). Any patient with a geocoding uncertainty above 8 had address data fed into the corrective geocoding algorithm (see section 4.3 on page 48). Patient meshblocks were then assigned for updated patient geocodes.

Only patients with a recorded National Health Index identifier (NHI) were included in the patient cohort. The NHI is an alphanumeric code assigned to individuals to uniquely identify them in the New Zealand health system. It is assigned to all new born babies, and to anyone who does not have one on their first contact with the health system. Most New Zealanders now have an NHI (MidCentral District Health Board, 2012). Some have more than one NHI, but these are linked together in the Ministry of Health master NHI database.

The NHI is necessary for this study as it is used to key-code and link data sets between primary and secondary care. Because the NHI is a unique identifier I have key-coded all NHIs using a technique where the alphanumeric representation of the NHI is converted to a purely numeric representation. This provides a layer of obfuscation which

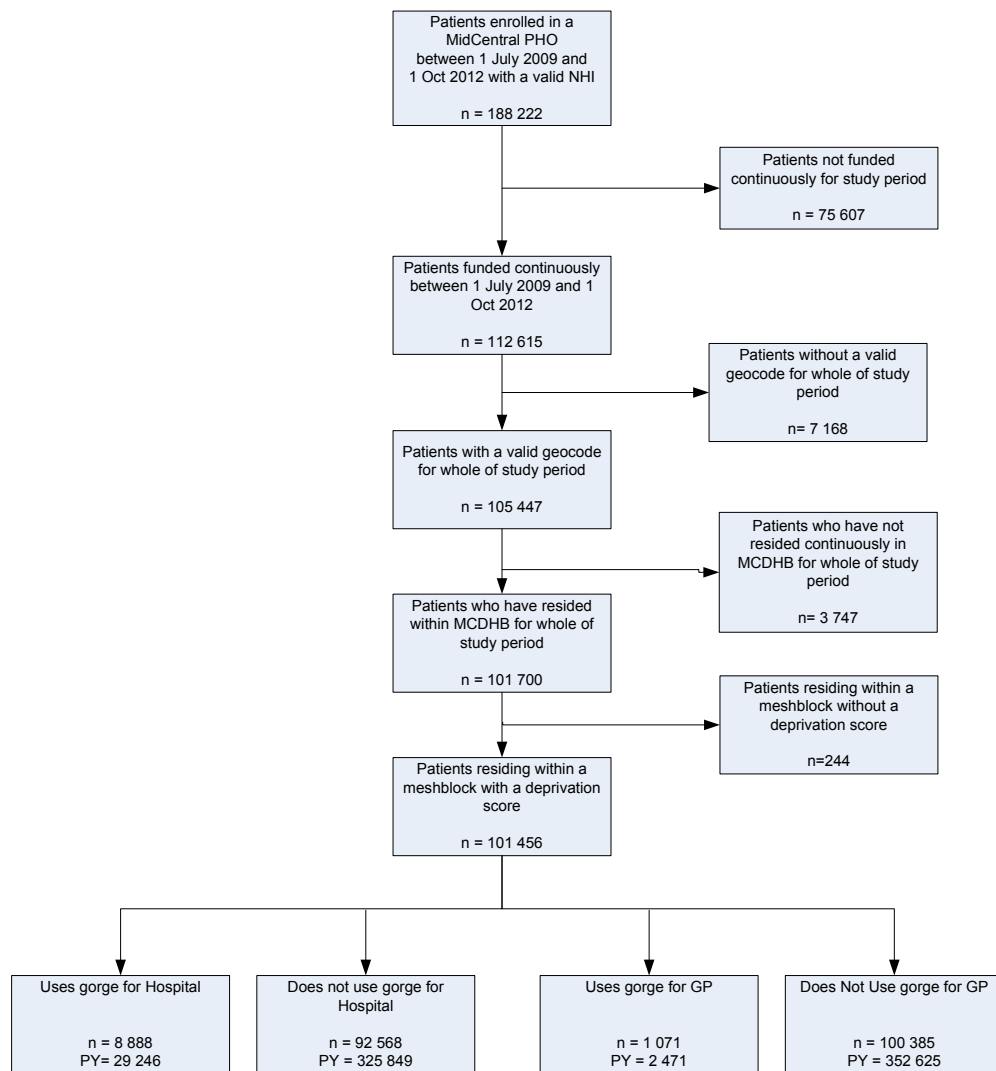


Figure 5.1: Cohort Selection Process with sample size (n) and person years (PY)

acts to protect people's identity as well as improving database match performance.

All patients enrolled in Central PHO between 1 July 2009 and 1 October 2012 were candidates for inclusion in the study. The study period consists of 14 three month PHO register periods. Those who were not funded continuously over the study period were excluded totaling 75,607 patients. Those who did not have an address that was geocoded

for each of the 14 funding periods were then excluded, totaling 7,168 patients. Those who did not reside in the MidCentral district for all 14 funding periods were excluded, totaling 3,747 patients. Those patients who resided in a meshblock without a deprivation score in any of the 14 periods were excluded, totaling 244 patients.

This resulted in a final eligible patient cohort of 101,456 patients, 53.9% of the PHO patient population. Patients were then assigned into four groups. One group represented those who had access to the hospital affected by the Manawatu Gorge closure (hospital-intervention group) and one who's hospital access was unaffected by the closure (hospital-control group). Another group was for those patients who's access to general practice was affected by the closure (GP-intervention group) and one for those who's GP access was unaffected (GP-control group). Each patient was assigned into one hospital and one GP group. When calculations were done on access to hospital services patients were analysed by hospital groups. When calculations were done for general practice services they were analysed by GP groups. Figure 5.1 shows the final counts for each group along with the number of person years of observations.

5.2 Potential Accessibility

5.2.1 Provider-to-Population Ratio

The provider-to-population ratio is formed by dividing two numbers that represent units of supply and demand. This unit may vary depending on what is being measured.

The provider based unit I used was mean provider days per quarter. To calculate this I used PHO Utilisation data. For each quarter I determined the number of days each provider had at least one utilisation record. Public Holidays and weekends were excluded from the calculation because general practices do not usually operate on these days. I

then divided the total sum of provider days in a quarter by the total number of available days.

I determined the population from the count of patients from PHO registers for each quarter. The PHO register represents all funded patients in the PHO.

To calculate the Provider-to-Population ratio I divided the number of providers by the total population each quarter. I also divided the population by the number of providers to give the inverse ratio.

5.2.2 Catchments

I undertook a catchment sensitivity analysis for both hospital and general practice services. Catchments were based on road network travel time using the MacRae travel time model (page 69). Calculations were made for each quarter of the study period. Calculations were performed only on patients with valid geocodes.

Travel times were rounded to the nearest minute. PostgreSQL was used to perform a query across each quarter's patient register to cumulatively count the number of patients within each minute's travel time of general practice and hospital.

The percentage of the population included in each travel time catchment was determined by dividing each catchment's cumulative patient count by the total patients in the register with a valid geocode.

5.2.3 Two-Step Floating Catchment Areas

The two-step floating catchment area (2SFCA) method uses the relationships between provider catchments, provider units and population catchments to calculate an index of accessibility.

I calculated 2SFCA accessibility indexes for general practice both while the gorge was open and closed. The gorge was open for a period of nine

quarters and closed or partially closed for a period of five. I used the patients in the PHO registers with a valid geocode.

I used 30 minutes road network travel time using the MacRae travel time model (page 69) for both provider and population catchments.

I performed population calculations using meshblocks. Travel time calculations were made from the population weighted centroids of each meshblock. The population used to calculate weighted geometric centroids was the funded population across each period. The population weight was the mean population across the period.

Provider catchments were calculated from the point location of general practices. Each general practice was assigned a provider weighting determined by the mean number of providers that worked at each practice per day. This was calculated from PHO utilisation data (see section 5.2.1 for implementation details).

A provider-to-population ratio was determined for each general practice at the 30 minute catchment threshold for the gorge open and closed periods. Each meshblock's accessibility for the two periods was then calculated by aggregating the provider-to-population ratios with the meshblock 30 minute catchment. These were categorised into deciles.

For each meshblock, I calculated the proportion index change from the gorge open period to the gorge closed period. I categorised each meshblock into one of three categories; a decrease by 20%; an increase by 20%; and a change less than 20%.

Meshblock accessibility index deciles and proportion change categories were visualised in cartographs using Quantum GIS.

5.2.4 Two-Step Floating Catchments with Distance Decay

Calculations were made using the 2SFCA method as described in section 5.2.3 but with a decay factor applied to all patients. All patients in the PHO registers with a valid geocode were included.

The decay factor (D_f) I used was 1.5 and the distance weighting function was a proportion of the distance of the population (d_p) from the catchment boundary (d_c). A minimum threshold (d_m) of 1 minute was used. Any travel time below this was assigned the maximal weighting of 1.

$$D = \left(\frac{d_c - d_p}{d_c - d_m} \right)^{D_f}$$

This was based on the method described by McGrail and Humphreys (2009).

5.2.5 Patient Choice

Patient choice of general practice was calculated for each meshblock in MidCentral for the quarter just prior to the gorge closure, and for the last quarter in which the gorge was closed.

For each meshblock I used a PostgreSQL query to identify practices with five or more patients registered and count them. This was done for the quarter 1 October 2011 and 1 October 2012, to represent the last period that the gorge was open and closed respectively. I created a cartograph showing the distribution of practice choice in Quantum GIS.

For each meshblock I calculated the difference between practice choice counts before and after the gorge was closed by subtracting the closed choice count from the open choice count. I created a cartograph showing the distribution of difference in practice choice by meshblock after the gorge was closed.

I plotted each meshblock by practice count and deprivation score using R. For each number of practices I plotted the maximum deprivation score of associated meshblocks and calculated a linear model.

5.3 Realised Accessibility

Palmerston North hospital supplied data from 1 July 2009 to 31 December 2012 of all ED attendance, Hospital admissions and Outpatient attendances. The data was identified by NHI and key-coded prior to removing from Palmerston North hospital. All data sets included either an attendance or admission date. Each data set contained specific information particular to the setting it represented.

ED data included an Australasian Triage Score. Hospital admission data included Length of Stay and Case Weighted Inlier Equivalence Separation values. Outpatient attendances included an attendance status code, which included a Did-Not-Attend flag which identified when patients didn't attend their appointment.

No hospital data sets contained demographic detail for patients. Demographic detail for all hospital data sets was derived by matching key coded identifiers to the PHO register data.

PHO registers were supplied by Compass Health on behalf of Central PHO, and contained a record for each patient funded in the PHO for each three month period. The demographic detail contained in the register file includes patient date of birth, gender, up to three ethnicity codes, patient addresses and geocoding information including the meshblock and socioeconomic deprivation.

General practice utilisation data was also supplied by Compass Health on behalf of Central PHO. This data was also identified by NHI and key coded. It contained the dates on which patients presented to general practice, a practice identifier and a registration number for the provider interacting with the patient.

5.3.1 Utilisation Rate Procedure

Utilisation rates were calculated for all measures in control and intervention groups for both gorge open and closed periods. All calculations

included data from 1 July 2009 to 31 December 2012 inclusive for patients in the cohort.

Utilisation counts were made for all groups by either day or month over the whole study period. Counts were converted to crude rates by using the cohort population for the quarter corresponding to the utilisation date as the denominator.

Crude rates were standardised. This was done using the direct method on age, deprivation, gender and ethnicity. Rates were standardised to the Central PHO register demographics.

I examined data to determine its likely distribution. All utilisation data met the criteria for a Poisson distribution. Poisson confidence intervals were calculated using the exact method at a 95% level.

Case Weighted Discharge data did not meet the criteria for a Poisson distribution. Its distribution shape suggested a log-normal distribution. Case Weighted Discharge data was compared by transforming it to natural logarithms and performing t-tests.

5.3.2 ED Attendance

Crude ED attendance rates were calculated for each year of the study period to compare to ED utilisation rates found in the literature. These rates were not standardised.

ED attendance rates were calculated by both month and day using the method described in section 5.3.1. Monthly rates were used for travel-time category analysis and daily rates were used for intervention analysis. The same analysis was undertaken for ED attendances with a Triage Score of 5 only.

ED attendance rates were also calculated for those falling within 5 minute travel time bands to ED up to 60 minutes. Travel times for the appropriate gorge open and closed period were used to classify patients. These rates were standardised and aggregated to means and Poisson confidence intervals calculated.

5.3.3 Hospital Admissions

Hospital admissions were calculated by month and day using the method described in section 5.3.1. Monthly rates were used for travel-time category analysis and daily rates were used for intervention analysis.

Case weighted discharge and length of stay (LOS) data was analysed by control and intervention groups for the gorge open and closed period. Case weights and LOS were aggregated to means. Case weight data was transformed to natural logarithms and confidence intervals were calculated from the log-normal distribution. Student's t-tests were applied to the means of the log transformed control and intervention data.

5.3.4 Hospital Outpatients Attendances

Hospital outpatient attendance rates were calculated by month and day using the method described in section 5.3.1. Monthly rates were used for travel-time category analysis and daily rates were used for intervention analysis. Any data that indicated non-attendance were excluded from this analysis. Attendance data that fell on weekends or public holidays were excluded from the analysis.

The same rate analysis was done for appointments where people did not attend the appointment.

5.3.5 General Practice Utilisation

General Practice attendance rates were calculated by month and day using the method described in section 5.3.1. Monthly rates were used for travel-time category analysis and daily rates were used for intervention analysis.

Attendance data that fell on weekends or public holidays were excluded from the analysis.

Chapter VI

Results

6.1 Last-Best-Known Geocoding

For all patients funded during the study period there were 140,250 unique addresses of which 21,902 that did not have sufficient geocoding certainty to be assigned a meshblock based on their latest geocode. Of these 6,799 previously ineligible address' geocodes were corrected using the last-best-known geocoding process (LBKGP). This resulted in an additional 12,323 patients being eligible for inclusion in the cohort. These results are summarised in Table 6.1.

Measure	Un-geocoded Patient Domiciles		
	Original	Best Match	Improvement(%)
Addresses	21,902	15,110	31.0%
Patients	33,262	20,939	37.0%

Table 6.1: *Improvement of Un-geocoded Patient Domiciles Using a Last-Best-Known Geocode Method*

The result of applying this process to the records in Table 4.2 on page 50 would identify the record for January 2008 as the best geocoding match.

6.2 Corrective Geocoding

The corrective geocoding algorithm (CGA) coded 140 addresses of the 220 in the validation set. This represents 63.3% of the sample.

Table 6.2 shows the validation results for each matching method in the CGA. Methods 1,2 and 4 were the most accurate for identifying the correct road with positive predictive values (PPV) of 100. Methods 1 and 4 were very accurate in providing a location within 94 and 32 meters respectively compared to the Google API location for the corrected address. Method 2 had a relatively large error of 6.1 km. A large proportion of addresses were corrected using method 5 with a high PPV of 96. Additionally the locations were very accurate for correct addresses. Incorrectly coded addresses contributed substantial error in method 5 with a mean of 20.9km.

Method	Population		n	Correct		Incorrect		PPV
	N	%		n	error (m)	n	error (m)	
1	2418	16.0%	30	30	94	-	-	100
2	654	4.3%	10	10	6,121	-	-	100
3	617	4.1%	10	9	92	1	23,455	90
4	179	1.2%	10	10	32	-	-	100
5	3258	21.6%	50	48	465	2	20,899	96
6	765	5.1%	10	8	321	2	12,424	80
7	1254	8.3%	20	8	3,225	12	67,871	40
NA	5965	39.5%	80	-	-	-	-	-
Total	15,110	100.0%	220	123	328	17	53,209	55.6

Table 6.2: *Corrective Geocoding Algorithm performance by road matching methodology*

A total of 5,855 additional addresses were geocoded using the CGA by using methods 1,4 and 5. These codes had a PPV of 98.0. An additional 12,654 addresses were geocoding by using the LBKGP and CGA. This represents 57.7% of originally un-geocoded addresses.

6.3 *Developing a Travel Time Model*

6.3.1 Derived Measures

Intersection Density

A comparison of urban boundaries as defined by Statistics NZ (grey shaded areas) and areas with intersection densities above 2 (brown shaded areas) is shown in Figure 6.1.

The areas defined by Statistics NZ are much larger than the area with intersection densities of three or higher. Both areas cover the main urban areas of Palmerston North, Feilding, Woodville and Pahiatua. Not shown in the figure are the other main urban areas of MidCentral, including Foxton, Levin and Otaki. These all exhibited a similar pattern to that shown in Figure 6.1.

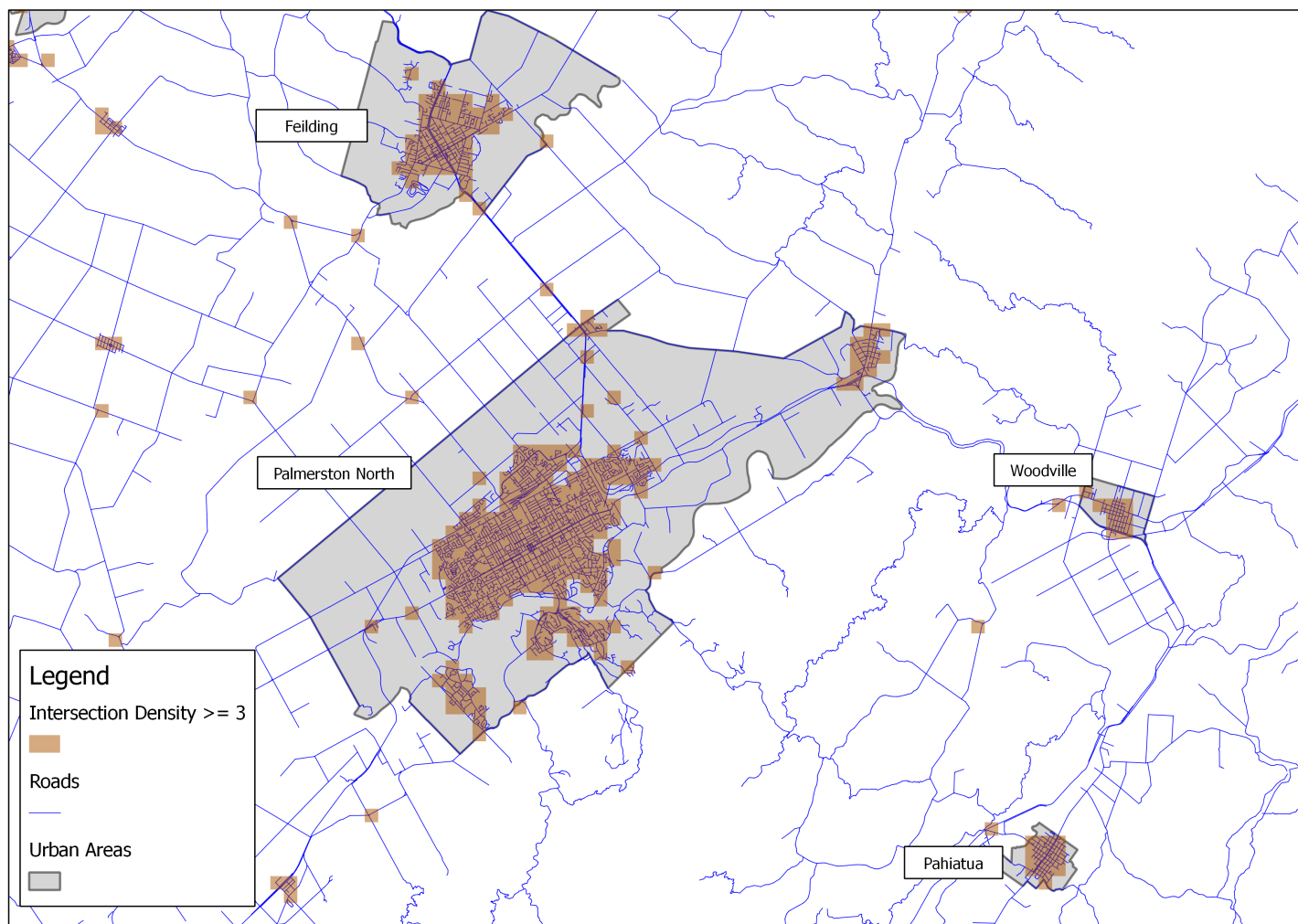


Figure 6.1: Comparison of Statistics NZ Urban Areas and Derived High Intersection Density Areas

Elevation

The interpolation of 20m contours produced results that were incredibly poor when compared to observed data through the Manawatu Gorge. Figure 6.2 shows the plot of predicted and observed data, along with a scatter plot and associated linear model with 95% confidence intervals.

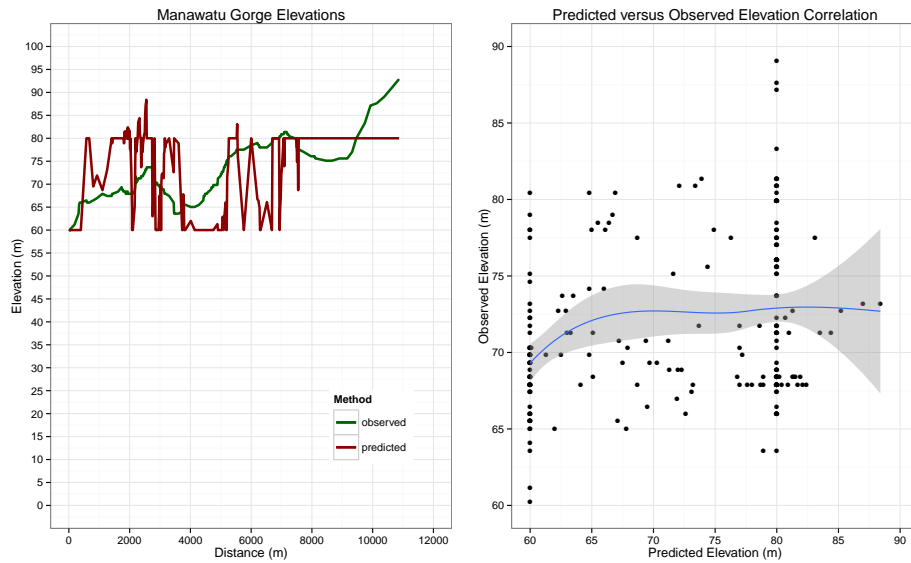


Figure 6.2: Performance of Interpolated Elevation Predictions Versus Observations

The plot of predicted and observed data by travel distance shows that the predicted data fluctuated between 60 and 80m. The observed data reveals a less erratic profile. The gorge appeared to climb from 60 to 95m over 11km. The journey is characterised by two hills peaking approximately 2km and 7km through the journey.

The scatter plot shows that there was very little correlation between the predicted and observed data. A large proportion of predictions were clustered around the 60 and 80m elevations. This was likely an artifact of using the 20m contour data for interpolation.

The elevation model performed poorly in the specific scenario of determining elevation of the Manawatu gorge from 20m contour data. The

r^2 value was only 0.08.

6.3.2 Speed & Attribute Correlation

The way in which attributes were correlated with observed speed is shown in Figure 6.3. These data represent all speed classes for simplicity of presentation. This grid was generated for each of the 7 speed classes of road available for analysis. Some categorical data tended to cluster travel speeds into narrow bands in the higher ranges. Road surfaces in class 3 had a range of limited speeds between 80 and 110 km/hr. Roads with an attribute indicating six lanes had a very narrow speed range of 80 to 110 km/hr.

The continuous data exhibited large clustering and large variance in lower ranges, but much narrower variance as the values became more extreme. Sinuosity Index had a high range close to 1 (0 to 110 km/hr); between 1.5 and 2.0 the range was smaller (5 to 60 km/hr); and greater than 2.0 it was the smallest (15 to 30 km/hr). A similar but less dramatic pattern was exhibited in Intersection Density. Total turning angle appeared to have a similar pattern, between 0 and 500 degrees, and then independently above 500 degrees. Maximum turning angle exhibited a similar pattern up to approximately 100 degrees.

The minimum travel speeds remained consistent across all variables, usually below 10 km/hr. The decrease in range of travel speeds was therefore mainly a function of the decrease in the maximum travel speed. The maximum values exhibited a roughly linear relationship.

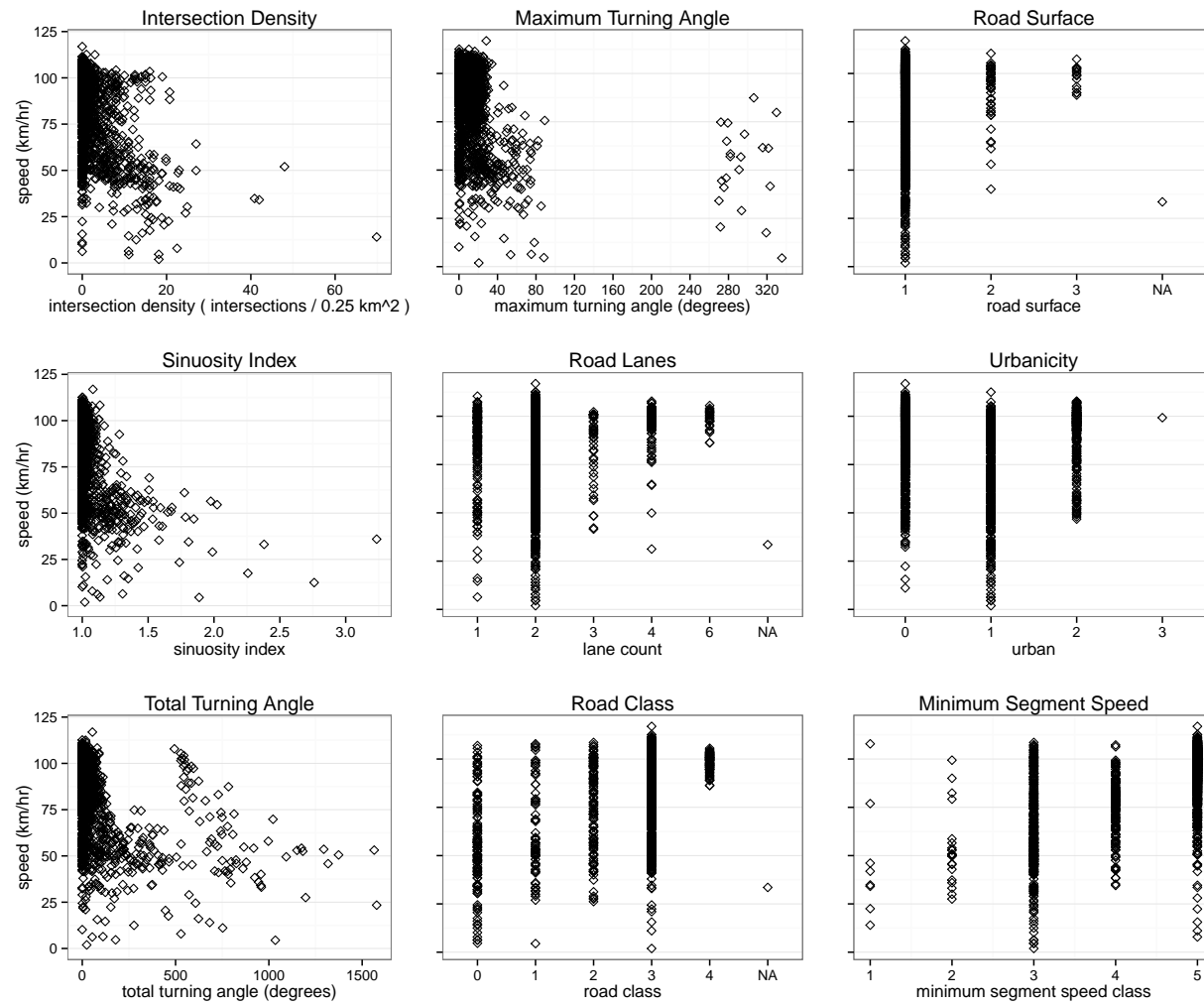


Figure 6.3: Matrix of Correlation Between Road Attributes and Speed Measurements

6.3.3 Linear Regression

The results of the AIC stepwise multiple linear regression can be seen in Table 6.3. The attributes with the most significant influence on travel speed were speed class, road class, sinuosity index, log of intersection density. The number of lanes was not significant at a 0.05 level.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	50.2	8.5	5.9	1.605E-08
Speed Class	12.1	1.2	10.4	6.3093E-21
Road Class	6.9	0.8	8.3	1.4968E-14
Sinuosity Index	-37.5	5.5	-6.8	1.2201E-10
log(Intersection Density)	-4.5	1.8	-2.4	0.016065
Lanes	1.7	1.0	1.7	0.091522

Table 6.3: AIC Stepwise Multiple Linear Regression of Road Attributes to Estimate Travel Speed

The adjusted r^2 of this model with observation data was 0.6667, showing a weak correlation. Informal testing of this model showed that results conformed poorly to test data providing extreme values for some trips.

6.3.4 Model Comparisons

Figure 6.4 shows how each model was correlated to each other. It reveals that all models were highly correlated with r^2 values ranging from 0.95 to 0.987. Overall the regression reveals that the Google model was approximately the same as the Basic model with some initial offset. Compared to the Google model the MacRae model was overall 20% slower and the Brabyn-Skelly model was 40% slower.

The MacRae and Basic model had the highest correlation with an r^2 of 0.987. The Google model compared to the Basic model had a coefficient of 1, with an intercept of 7.9.

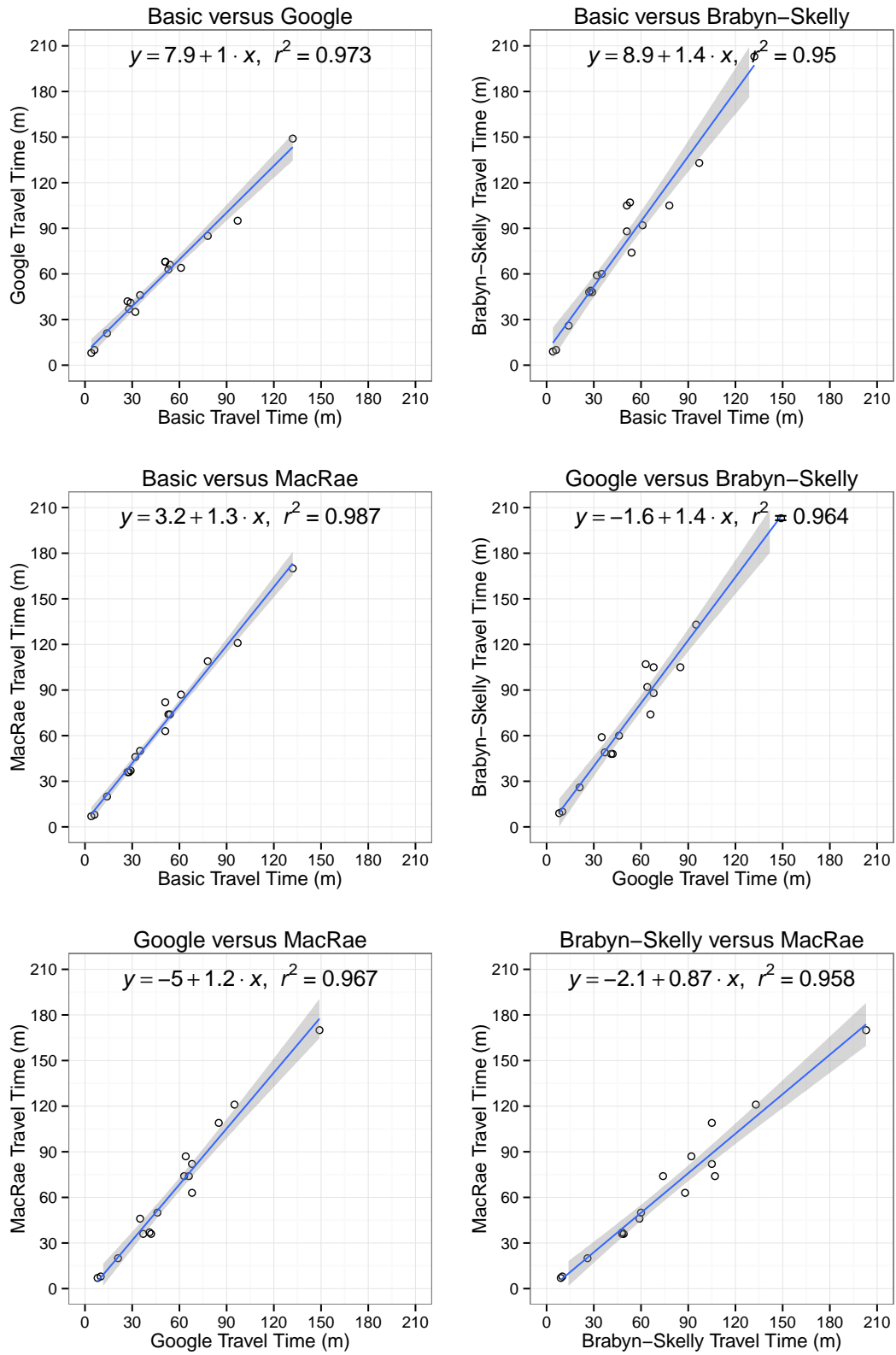


Figure 6.4: Travel Time Model Regression Comparisons

Figure 6.5 shows how the models compared against each other over different modeled journeys of varying distance measured against the Basic travel time. Excluding the Basic model which is included for reference only, the shortest trip ranged from 7 - 9 minutes while the longest ranged from 149 - 203 minutes.

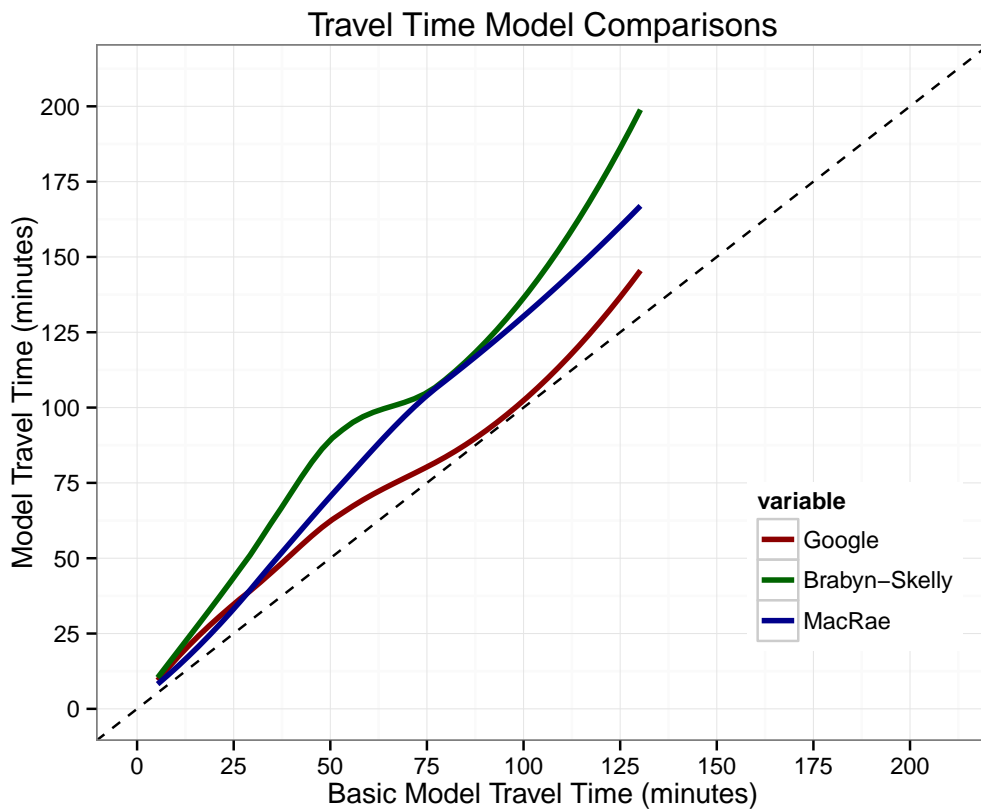


Figure 6.5: *Travel Time Model Predictions*

The Google model appeared to be the fastest and had travel times slower than the reference up to 60 minutes and over 90 minutes. Between these the Google Model approximated the Basic travel time. The two models varied substantially due to the intercept of the correlation, with a 7.9 minute offset.

The Brabyn-Skelly model was the slowest of all the models. For trips under 10 minutes it approximated the other model times. Between 55

and 80 minutes it approximated the MacRae model and beyond this it was consistently slower than both. At the most extreme distance measured, it was 54 minutes (36.2%) slower compared to the Google model.

The MacRae model approximated the Google model up to 55 minutes, and then became substantially slower through to 80 minutes. After this mark it performs faster than the Brabyn-Skelly model, and slower than the Google model.

6.3.5 Empirical Validation

Figure 6.6 shows the distribution of the observations for each survey site and travel direction with the predicted travel times from each model. The variances for each trip and direction of travel were different.

The Brabyn-Skelly model consistently over estimated journey times compared with observed median travel times for all trips except the Rimutaka Hill Road. The MacRae model underestimated travel times for trips across the Rimutaka Hill Road and Saddle Road. Google overestimated for the Pahiatua Track and Saddle Roads.

Figure 6.7 shows the distribution of the errors for each model. Google provided the least precise model having the greatest error range in both minutes and as a percentage of travel time. Google however represented the most accurate predicted travel time with a median error of -0.2 minutes.

The MacRae model was similarly accurate as Google (median error of -0.5 minutes) but was more precise with a lower range and inter-quartile range in both gross and percentage error.

The Brabyn-Skelly model was the least accurate and precise. While its median error was only 1.3 minutes its median percentage error was 15.9. Its precision was similar to the MacRae model and better than Google.

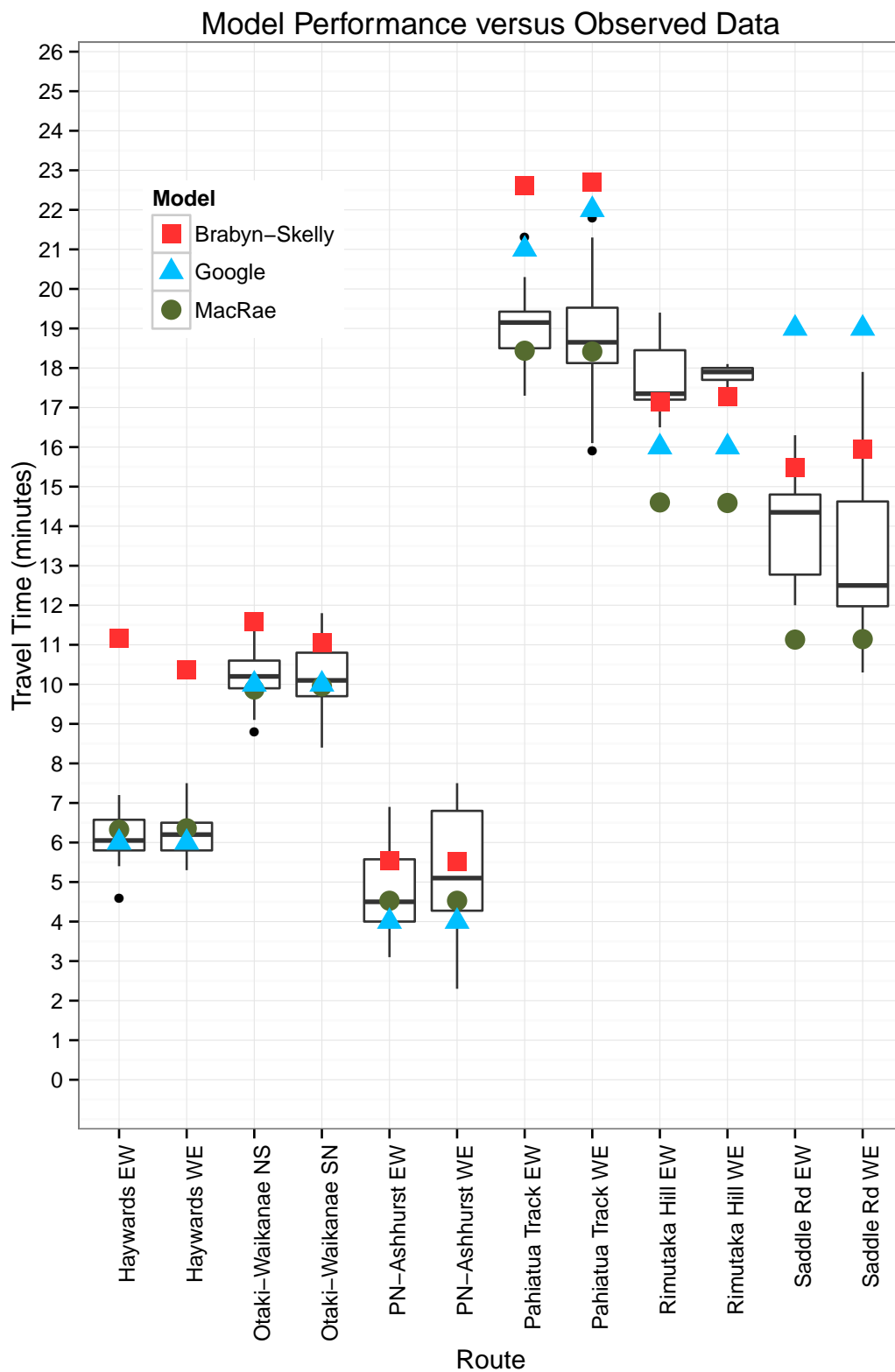


Figure 6.6: Travel Time Model Predictions Compared with Traffic Survey Data

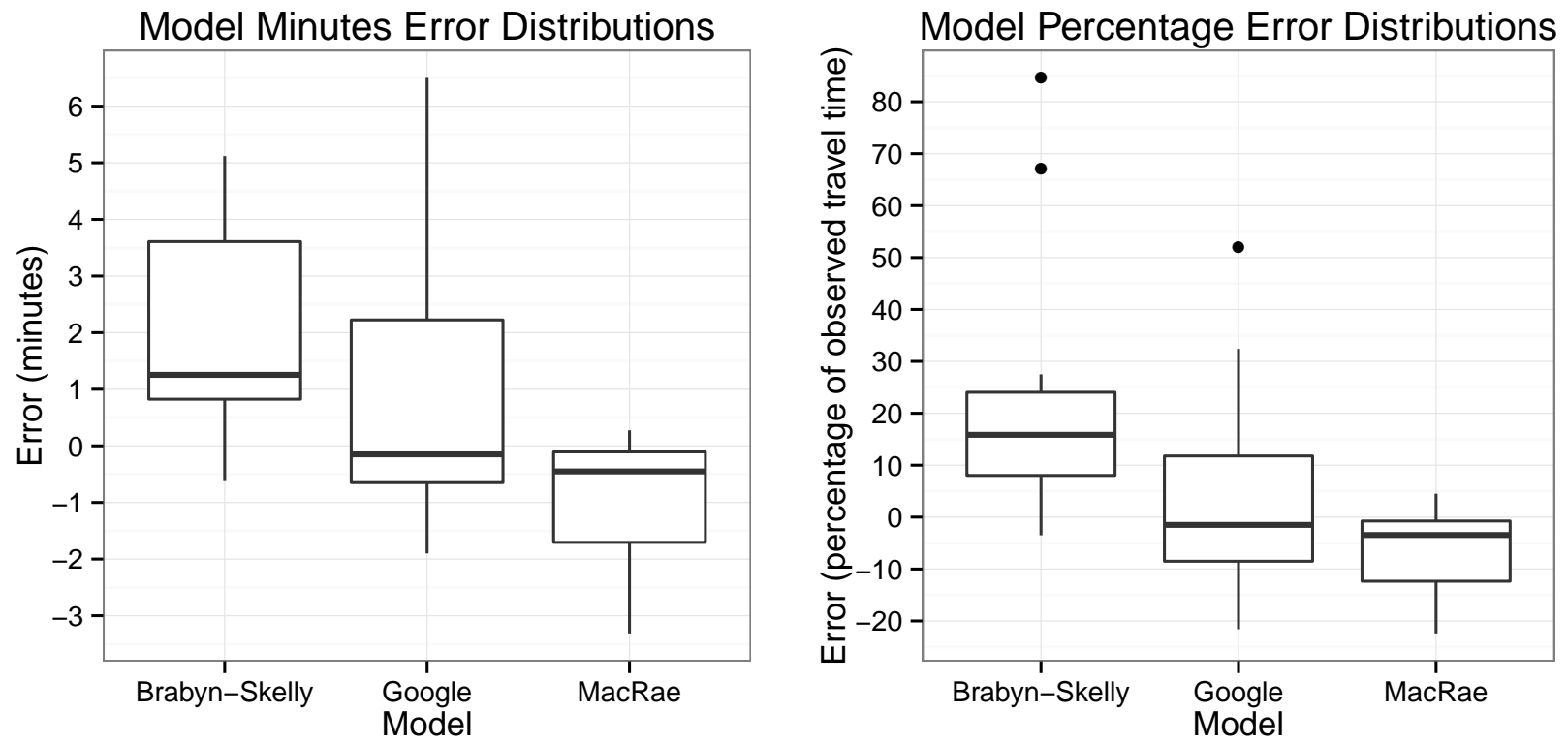


Figure 6.7: Travel Time Model Error Distributions

6.4 Cohort Demography

A total of 6,970 patient register records were excluded because they did not have a valid NHI. Because the NHI is the method by which patients are uniquely identified in the data set, it was not possible to definitively determine how many patients this was exactly. This number is an upper bound and is likely to be well less than the total number of patient records, as there were 14 funding periods in the study period and patients had a record once for each funding period they were in the PHO. This resulted in a lower bound estimate of patients excluded of 498 patients.

The patient cohort was constructed of 101,701 patients that met the inclusion criteria. The cohort was grouped into control and intervention groups. Table 6.4 shows the mean travel times to Palmerston North hospital and each patients funded general practice by cohort group.

Group	Mean Travel Time (minutes)	
	Hospital	GP
Control	20	7
Intervention	46	39

Table 6.4: Mean travel times to hospital and general practice by cohort group.

Figure 6.8 shows the demographics of those patients included in the study cohort compared with that of the PHO funded population. Age exhibited the most difference between cohort and funded population profiles of the four demographic categories.

People aged under 5 years, 16-38 years and those older than 89 years were excluded from the study in higher proportions than other age groups. The largest difference was those under the age of five. The proportion of under fives was three times higher in the funded population than the cohort. In general toddlers and young adults were

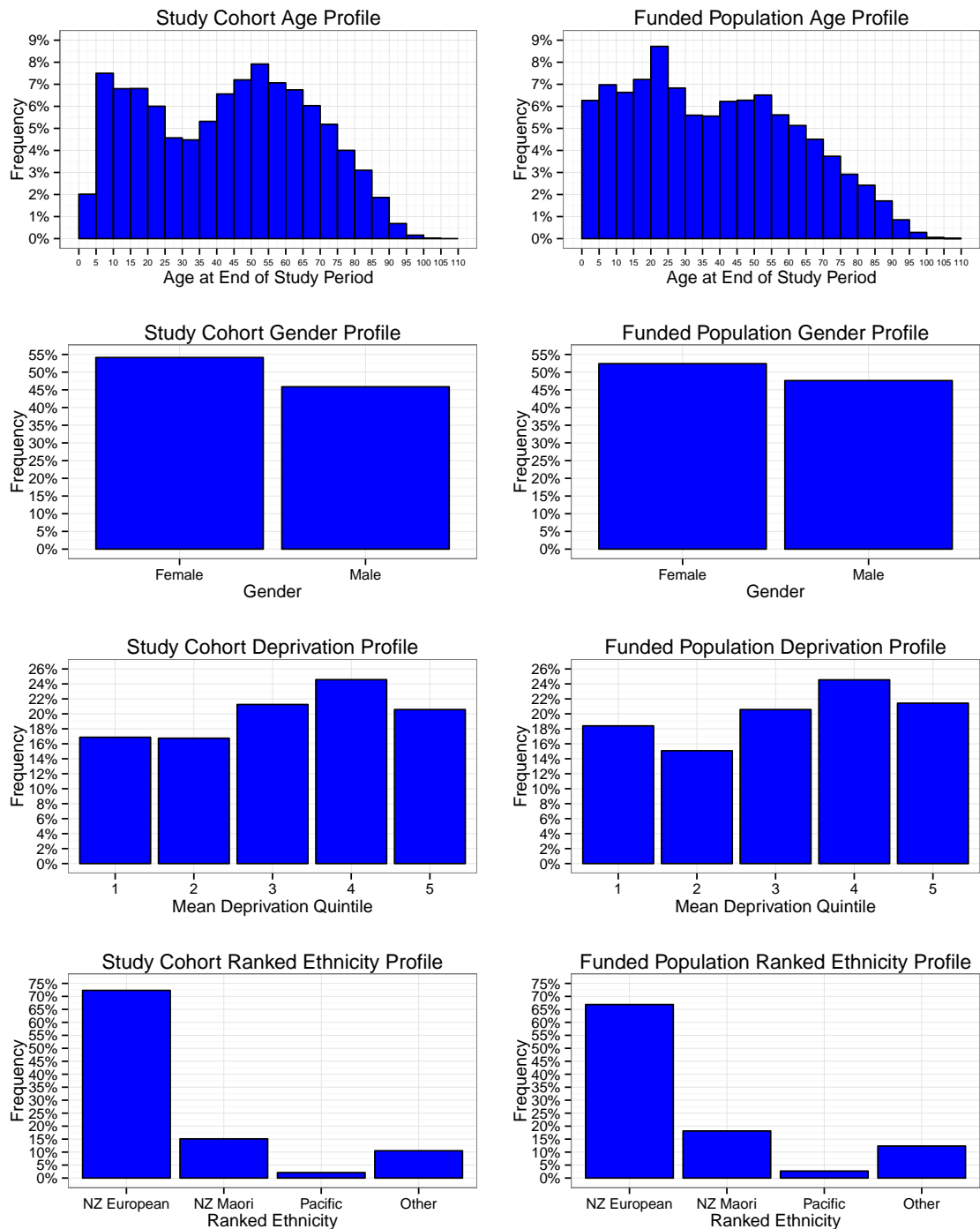


Figure 6.8: *Demographics of Study Cohort and Funded Population*

under-represented in the cohort, while middle aged adults are over-represented.

Deprivation was almost the same. There was some small variation in quintiles 1 and 2. This was only by a few percent however.

Figure 6.9 shows that the proportion of deprivation in the cohort was highly dependent on ethnicity. NZ European and Other ethnic groups had a relatively even distribution of deprivation. NZ Maori and Pacific both had a disproportionate number in higher deprivation groups.

Figure 6.10 shows how deprivation in the cohort was distributed by geography. Deprivation was highest in urban areas. All towns and cities had areas of substantial deprivation, almost exclusively containing the most deprived quintile. By contrast the least deprived areas were rural in close proximity to urban centres. Palmerston North and Feilding had areas of low deprivation encompassing them, while Dannevirke had an area of low deprivation to the east, and Pahiatua to the south-east.

The predominant ethnicity of the funded population and cohort was NZ European, with about 15% NZ Maori and 2.5% Pacific Islanders. The cohort had a slight over-representation of NZ European and slight under-representation of NZ Maori compared with the funded population.

Figure 6.11 shows how each ethnic group was distributed across the MidCentral district at a macro level, while figure 6.12 has the micro view. Each meshblock for each ethnic group is classified into a quartile for the proportion of the study cohort that resided there. New Zealand Europeans lived predominantly in the northern and western areas of the district and had higher concentrations in rural areas. There was a very low proportion of NZ Europeans that lived in Dannevirke. This is in stark contrast to the NZ Maori population, who had high rates of urban distribution and lower rates in the eastern areas of MidCentral. The Pacific population, although small, lived almost exclusively in urban areas. The eastern area of MidCentral is where Other ethnic groups had the highest rural presence.

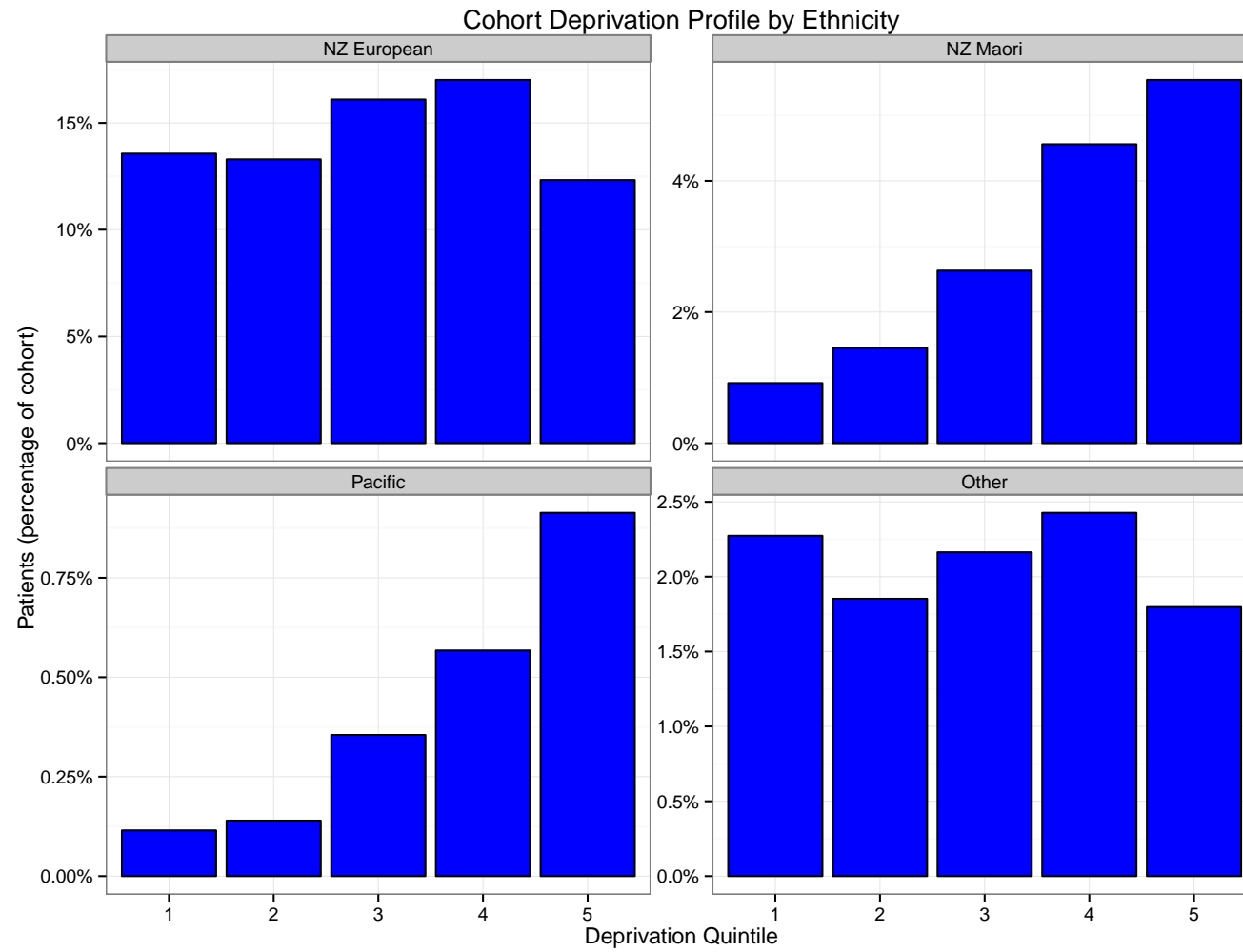


Figure 6.9: Deprivation Profile of Cohort Patients by Ranked Ethnic Groups

Geographic Distribution of Deprivation in MidCentral

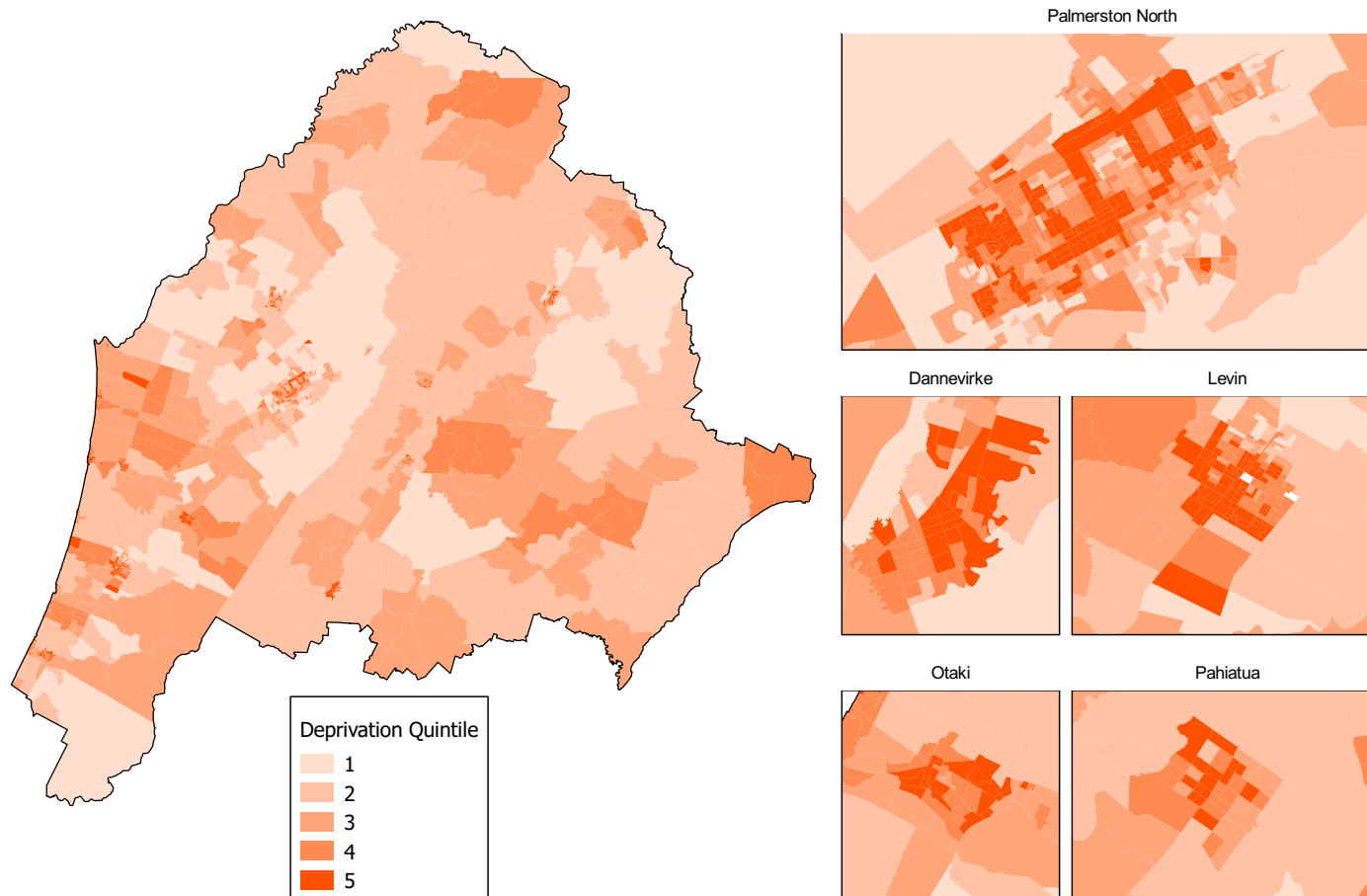


Figure 6.10: Deprivation Distribution by Meshblock

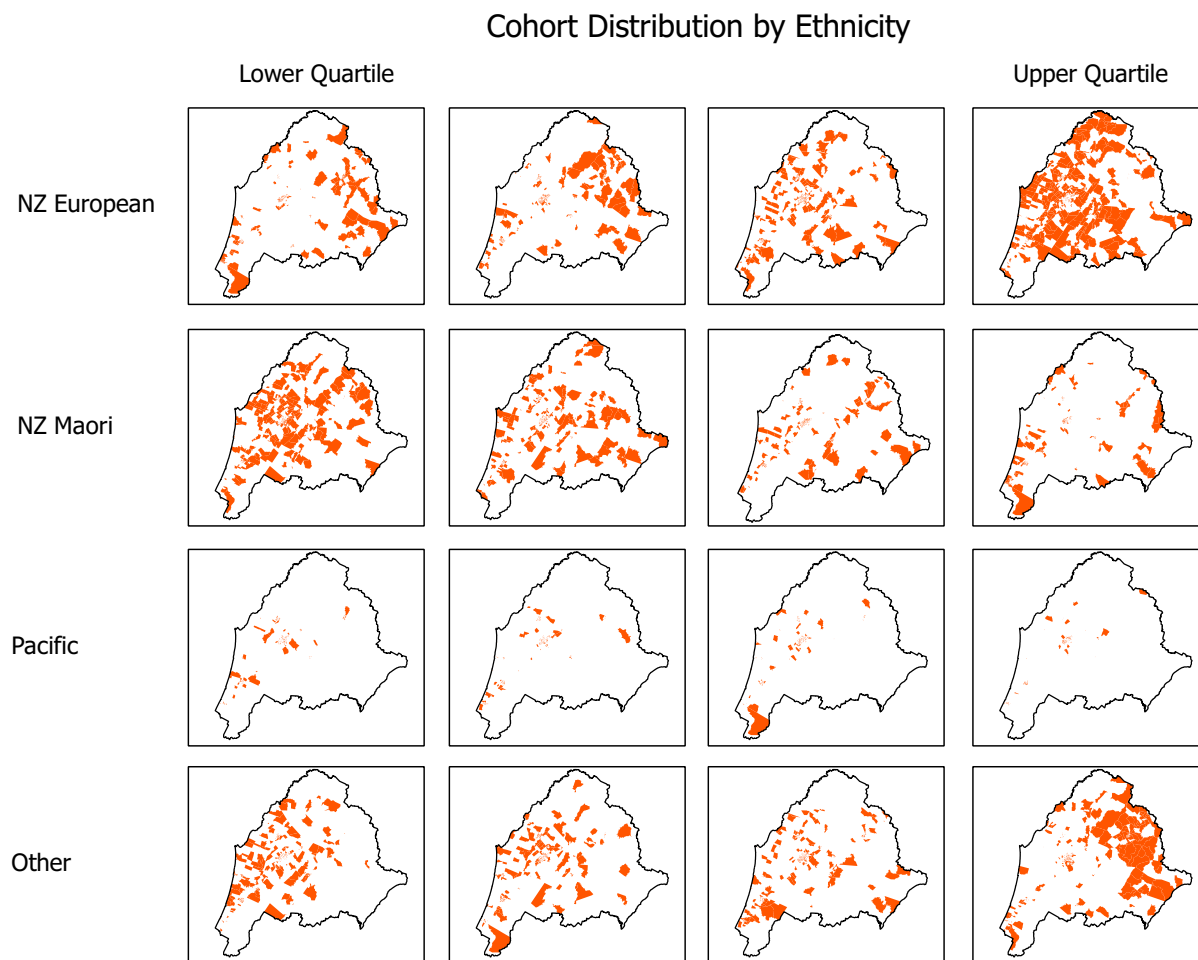


Figure 6.11: *Ethnicity Proportional Distribution Matrix*

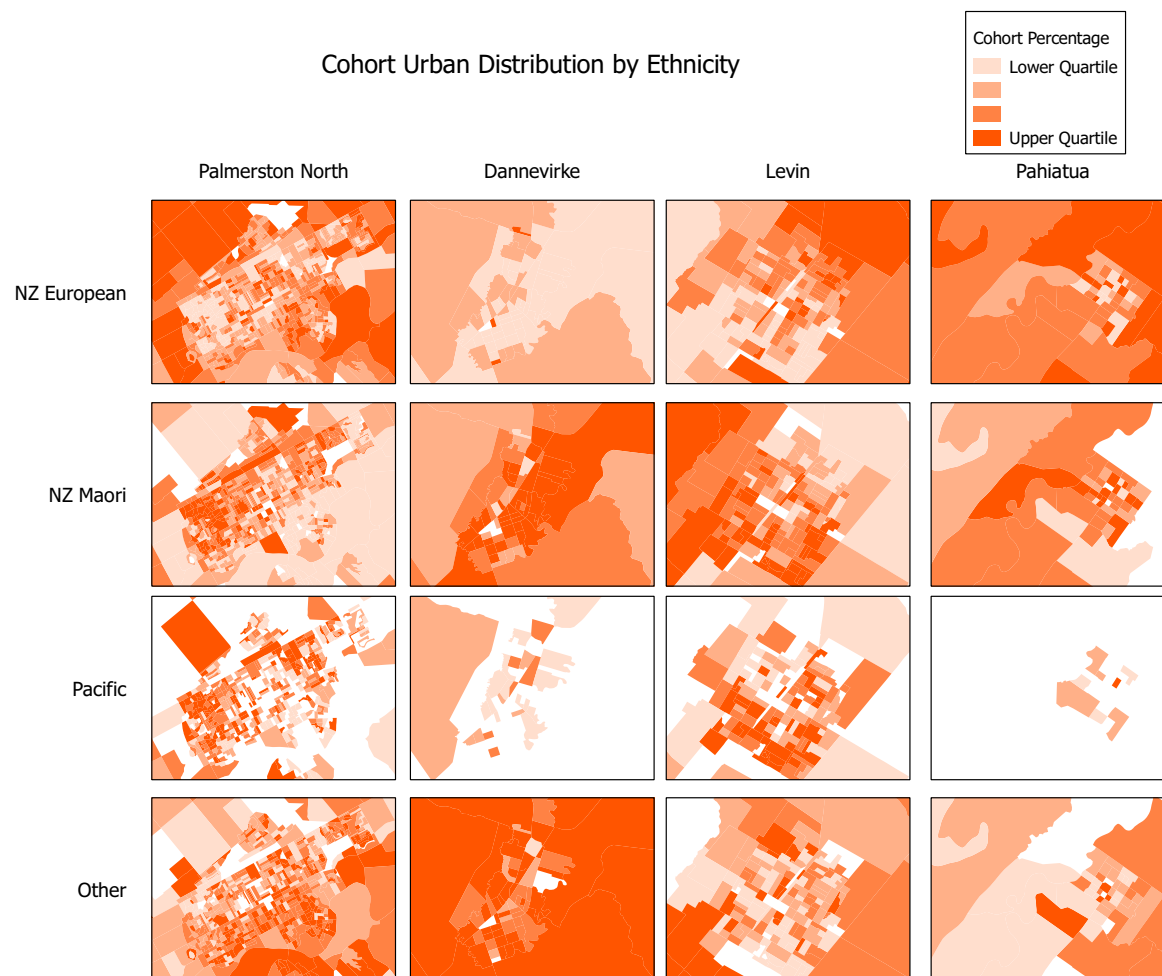


Figure 6.12: *Ethnicity Distribution in Urban Areas*

Figure 6.13 shows the cohort age profile for each ethnic group. The NZ Maori and Pacific age profiles showed an almost linear trend with high proportions of their population under 20 and a steady decline in each age group after that. The NZ European group had two distinct peaks, with the first around 20-25 and the second around 50-55. There was a steady decline in the proportion of those older than 55 in the NZ European group. The Other group exhibited a single flat long peak with a steady decline in population proportion after the age of 45.

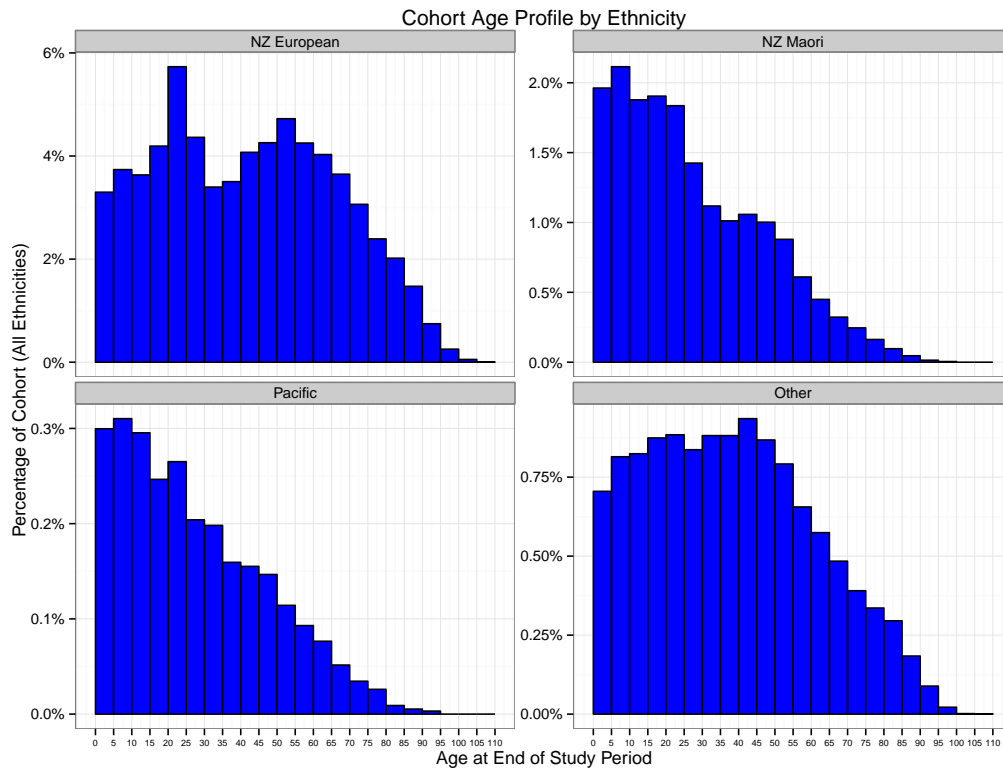


Figure 6.13: Age Profile of Study Cohort by Ethnicity

Emergency Department, Hospital Admissions and Outpatients all exhibited increasing utilisation rates from the least deprived to the most deprived groups.

6.5 Potential Accessibility

6.5.1 Travel Time Isochrones

Isochrones to Palmerston North Hospital while the gorge was open are shown in Figure 6.14 and Figure 6.15 in five minute intervals. Palmerston North's urban area had three zones of travel times varying from just under 15 minutes in the most southwestern suburbs, to under 10 minutes in the southern, western and eastern suburbs, and under 5 minutes in the central suburbs immediately around the hospital.

Four corridors of extended travel distance for each isochrone exist. These follow the directions in the south-west, north-north-east, north-east-east and south-east.

MidCentral's east had more area with greater than 60 minutes travel time to the hospital. The northern and southern fringes of the district had significant areas also. The western areas were almost exclusively within the 60 minute envelope.

The area affected by increased travel time to Palmerston North hospital is also shown in Figure 6.14. Dannevirke and Woodville were both affected, while Pahiatua was on the cusp of the affected area. At 2,950 km² the affected area represented 33.3% of the MidCentral District.

6.5.2 Provider-to-Population Ratio

Table 6.5 shows accessibility in the MidCentral district for general practice services over the study period using a provider-to-population ratio.

The patient-to-provider rate decreased over this time, suggesting there was an increase in accessibility across the MidCentral district. The total number of patients funded in Central PHO increased over the study period by 1.4%. The total number of providers also increased but by a higher rate of 5.8%. Because the rate of increase of providers was higher than patients the overall accessibility increased by 4.1%.

Travel Time Isochrones to PN Hospital

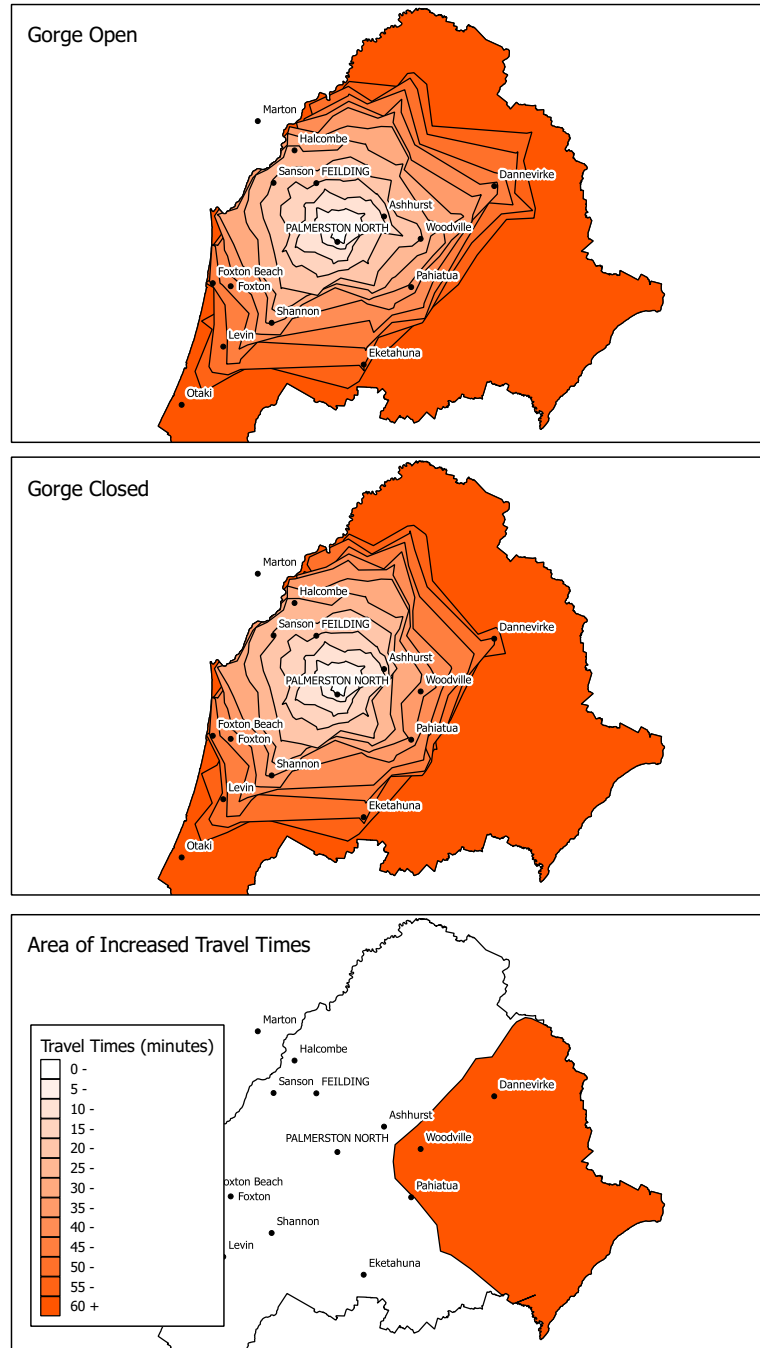


Figure 6.14: Travel Time Isochrones from Palmerston North Hospital with Gorge Open and Closed

Urban Travel Time Isochrones to PN Hospital

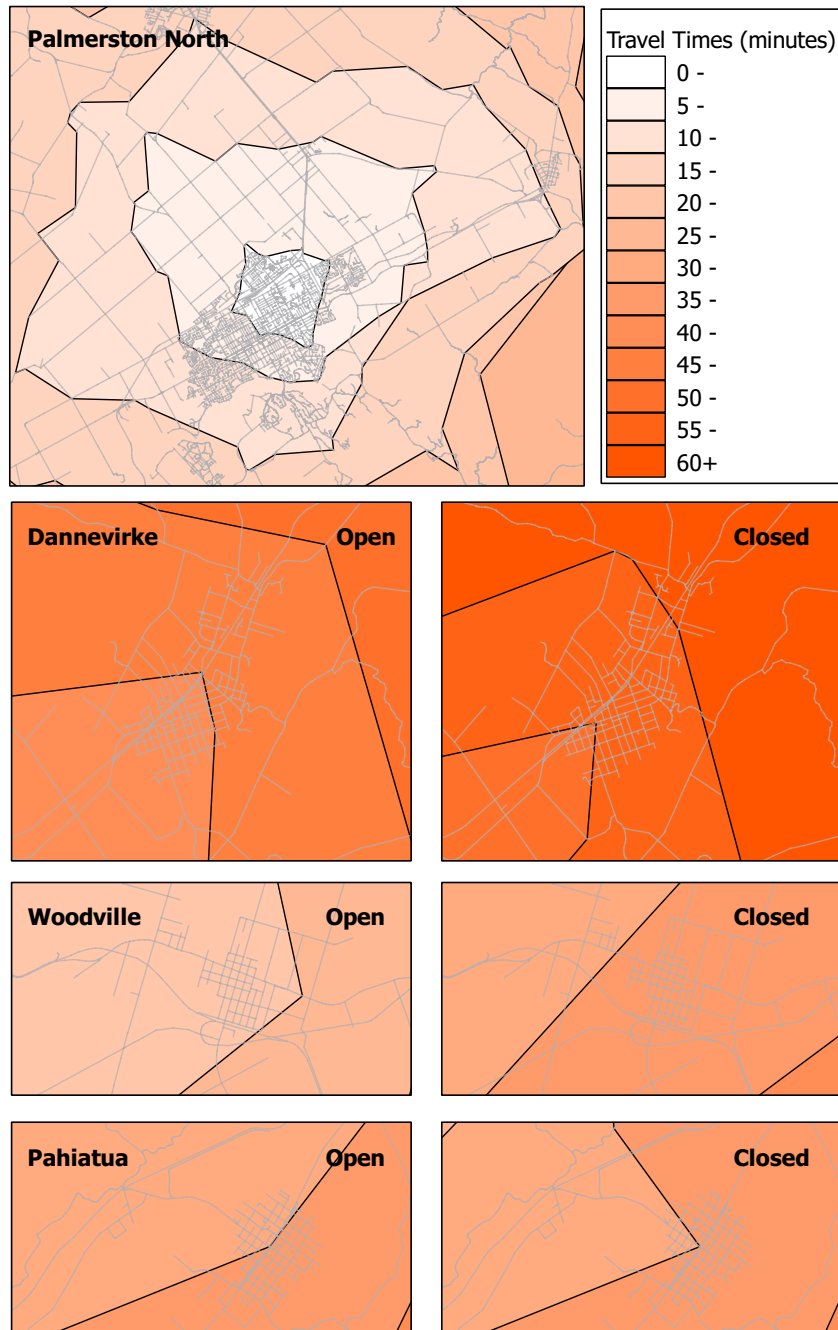


Figure 6.15: Urban Travel Time Isochrones from Palmerston North Hospital with Gorge Open and Closed

Quarter	Patients	Providers	Patients per Provider	Providers per Patient
1-Jul-09	150,710	100.4	1,501	6.7×10^{-04}
1-Oct-09	150,640	101.6	1,483	6.7×10^{-04}
1-Jan-10	150,627	103.1	1,461	6.8×10^{-04}
1-Apr-10	150,993	101.7	1,485	6.7×10^{-04}
1-Jul-10	151,032	104.7	1,443	6.9×10^{-04}
1-Oct-10	151,126	109.7	1,378	7.3×10^{-04}
1-Jan-11	151,255	111.8	1,353	7.4×10^{-04}
1-Apr-11	152,228	114.7	1,327	7.5×10^{-04}
1-Jul-11	152,644	112.0	1,363	7.3×10^{-04}
1-Oct-11	153,024	110.0	1,391	7.2×10^{-04}
1-Jan-12	152,956	108.1	1,415	7.1×10^{-04}
1-Apr-12	152,690	108.5	1,407	7.1×10^{-04}
1-Jul-12	152,693	107.5	1,420	7.0×10^{-04}
1 Oct-12	152,792	106.2	1,439	7.0×10^{-04}

Table 6.5: *Provider-to-Patient Ratio for General Practitioners in MidCentral*

Accessibility reached a peak in the quarter starting April 2011. After this period accessibility began to decline steadily, but at a rate slower than what it grew. During the gorge closure, accessibility had already begun to decline.

6.5.3 Catchments

Access to general practice by travel time catchments can be seen in Table 6.6. In July and October 2009, over 95% of the population were within 30 minutes travel time of general practice, but from January 2010 onwards, there was a large jump to 40 minutes to reach the same 95% coverage. During the closure of the Manawatu gorge the travel time needed to reach the 95% threshold didn't change, but the coverage at this level dropped by 0.2%. At 30 minutes travel time, the drop during the gorge closure was 0.3%.

Quarter	Travel Time to GP (minutes)												
	28	29	30	31	32	33	34	35	36	37	38	39	40
1-Jul-2009	94.9	95.4	95.6	95.9	96.2	96.4	96.6	96.8	96.9	97.1	97.2	97.4	97.5
1-Oct-2009	94.9	95.3	95.5	95.8	96.1	96.4	96.6	96.8	96.9	97.1	97.2	97.3	97.5
1-Jan-2010	92.2	92.7	93.0	93.3	93.6	93.9	94.1	94.3	94.4	94.6	94.7	94.9	95.2
1-Apr-2010	92.3	92.8	93.0	93.3	93.6	93.9	94.1	94.3	94.4	94.6	94.7	94.9	95.2
1-Jul-2010	92.2	92.7	93.0	93.2	93.6	93.8	94.0	94.2	94.4	94.5	94.7	94.8	95.2
1-Oct-2010	92.2	92.7	92.9	93.2	93.6	93.9	94.1	94.2	94.4	94.6	94.7	94.9	95.2
1-Jan-2011	92.2	92.7	92.9	93.2	93.6	93.9	94.1	94.2	94.4	94.6	94.7	94.9	95.2
1-Apr-2011	92.2	92.7	93.0	93.3	93.6	93.9	94.1	94.3	94.4	94.6	94.7	94.9	95.2
1-Jul-2011	92.2	92.7	93.0	93.3	93.6	93.9	94.1	94.3	94.5	94.6	94.8	94.9	95.3
1-Oct-2011	91.9	92.4	92.7	93.0	93.3	93.6	93.8	94.0	94.2	94.4	94.5	94.7	95.1
1-Jan-2012	91.9	92.4	92.7	93.0	93.3	93.6	93.8	94.0	94.2	94.4	94.5	94.7	95.1
1-Apr-2012	91.9	92.4	92.7	93.0	93.3	93.6	93.8	94.0	94.2	94.4	94.5	94.7	95.1
1-Jul-2012	91.9	92.4	92.7	92.9	93.3	93.5	93.8	94.0	94.2	94.3	94.5	94.7	95.1
1-Oct-2012	92.2	92.7	93.0	93.3	93.6	93.9	94.1	94.3	94.4	94.6	94.7	94.9	95.2

Table 6.6: Percentage of population within travel time to General Practice (shaded areas show travels times not meeting catchment threshold of 95%)

Access to after hours care through ED can be seen in and Table 6.7. Just over 93% of the population were within the 60 minute travel time catchment for after-hours care in normal circumstances. An additional 7 minutes were needed to bring this to 95%.

The closure of the Manawatu gorge had a small impact on access to ED. At sixty-minutes the population coverage dropped by 1.2%, and required 68 minutes to reach the 95% threshold.

6.5.4 Two-Step-Floating Catchments

Figure 6.16 shows two-step floating catchment accessibility indices for MidCentral during the period while the Manawatu Gorge was open. The index categories are divided into deciles.

The western aspects of MidCentral were well serviced, particularly around Palmerston North, while the northern and eastern areas were not. Areas around Dannevirke in the north showed moderate accessibility.

The last decile has a very large accessibility index range from 2.7660 - 838.8810.

Figure 6.17 shows the percentage difference in accessibility indices for each meshblock during the gorge closure. Differences are shown in three groups; those meshblocks who's indices either increased or decreased by 20% and those where the change was less than 20% in either direction. The twenty percent threshold was chosen arbitrarily.

There were few areas which experienced more than 20% increase during the gorge closure. The majority of meshblocks remained relatively unchanged. A substantial number experienced less access predominantly clustered in the north-west and in and around Dannevirke, Woodville and Pahiatua.

Quarter	Travel Time to Hospital (minutes)										
	60	61	62	63	64	65	66	67	68	69	70
1-Jul-2009	93.3	93.5	93.9	94.3	94.4	94.7	94.9	95.5	95.9	96.3	97.0
1-Oct-2009	93.3	93.6	93.9	94.3	94.4	94.7	94.9	95.5	95.9	96.3	97.1
1-Jan-2010	93.3	93.6	93.9	94.3	94.4	94.7	94.9	95.5	95.9	96.4	97.1
1-Apr-2010	93.3	93.6	93.9	94.3	94.5	94.7	94.9	95.5	95.9	96.4	97.1
1-Jul-2010	93.3	93.6	93.9	94.3	94.4	94.7	94.9	95.5	95.9	96.4	97.1
1-Oct-2010	93.3	93.6	93.9	94.3	94.4	94.7	94.9	95.5	95.9	96.4	97.0
1-Jan-2011	93.3	93.6	93.9	94.3	94.5	94.7	94.9	95.5	95.9	96.4	97.0
1-Apr-2011	93.3	93.6	93.9	94.3	94.5	94.7	94.9	95.5	95.9	96.4	97.0
1-Jul-2011	93.3	93.6	93.9	94.3	94.5	94.7	94.9	95.5	95.9	96.4	97.1
1-Oct-2011	92.1	92.9	93.3	93.7	93.9	94.1	94.3	94.9	95.3	95.8	96.5
1-Jan-2012	92.1	92.9	93.3	93.7	93.9	94.1	94.3	94.9	95.3	95.8	96.5
1-Apr-2012	92.1	92.9	93.3	93.7	93.9	94.1	94.3	94.9	95.4	95.8	96.5
1-Jul-2012	92.2	93.0	93.3	93.7	93.9	94.2	94.4	94.9	95.4	95.8	96.5
1-Oct-2012	93.4	93.7	94.0	94.4	94.6	94.8	95.0	95.6	96.0	96.5	97.1

Table 6.7: Percentage of population within travel time of Palmerston North Hospital (shaded areas show travels times not meeting catchment threshold of 95%)

General Practitioner Access in MidCentral using Two-Step Floating Catchment Method

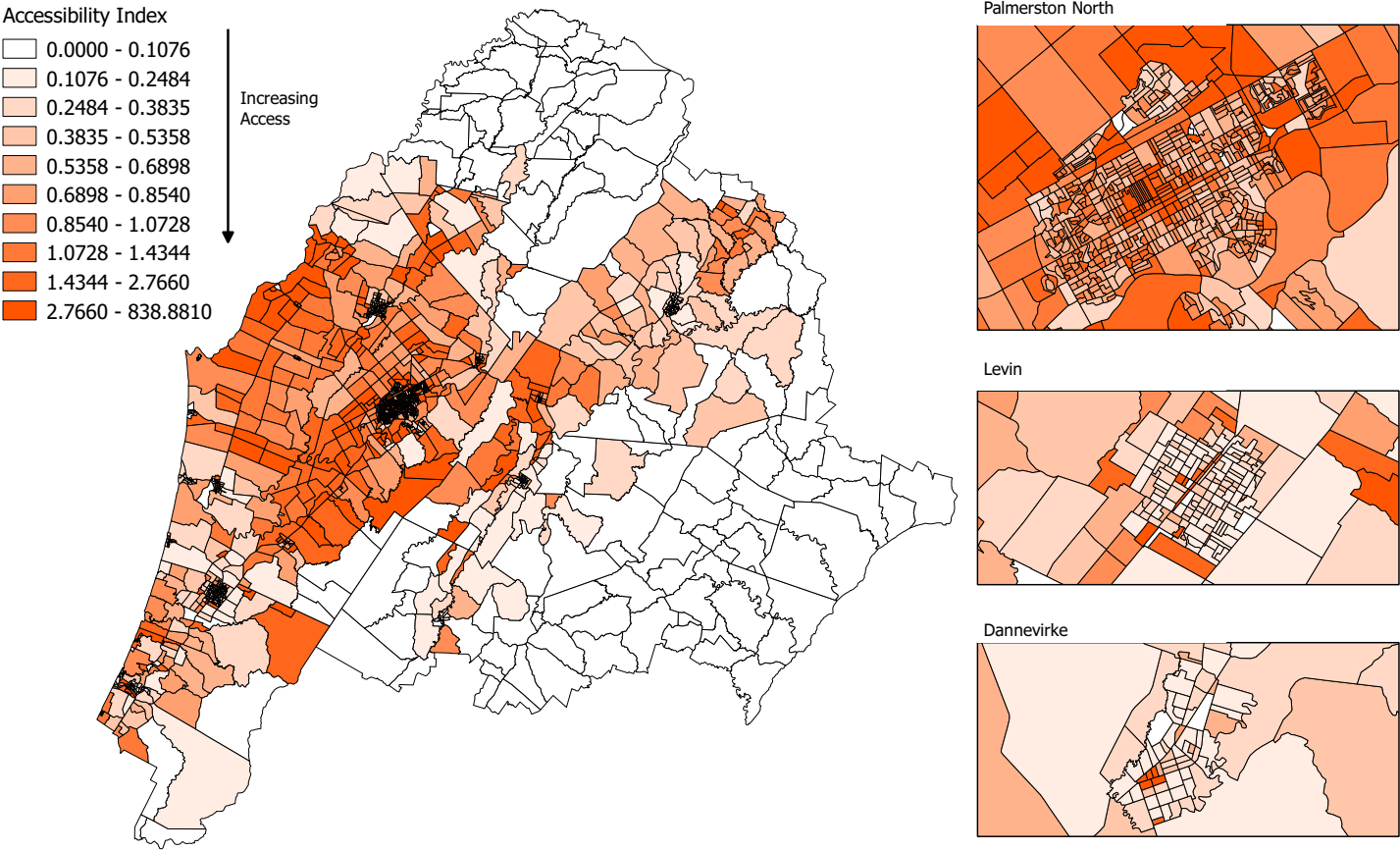


Figure 6.16: Accessibility for General Practices Using Two Step Floating Catchment Method With Gorge Open

General Practitioner Two Step Floating Catchment Ratio Differences Gorge Open versus Closed

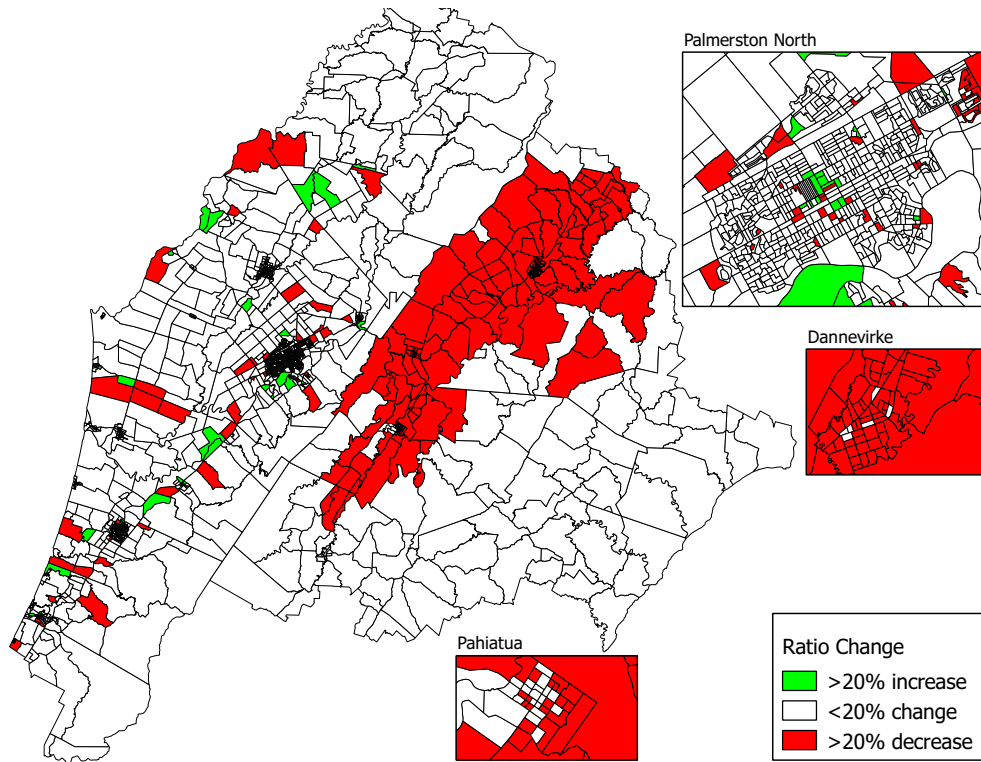


Figure 6.17: Accessibility Change for General Practice using Two Step Floating Catchment After Gorge Closed

6.5.5 Two-Step-Floating Catchments With Distance Decay

Figure 6.18 shows the accessibility indices for each meshblock calculated with a 2SFCD method. The pattern of accessibility was roughly similar to the 2SFCA. The north and far eastern areas showed very low accessibility, while the region around Dannevirke, Woodville and Pahiatua showed moderate access indexes. When compared to the 2SFCA the area around Dannevirke had lower accessibility.

The areas with the least accessibility were very similar to those seen in the 2SFCA.

General Practitioner Access in MidCentral using Two-Step Floating Catchment Method
with Distance Decay

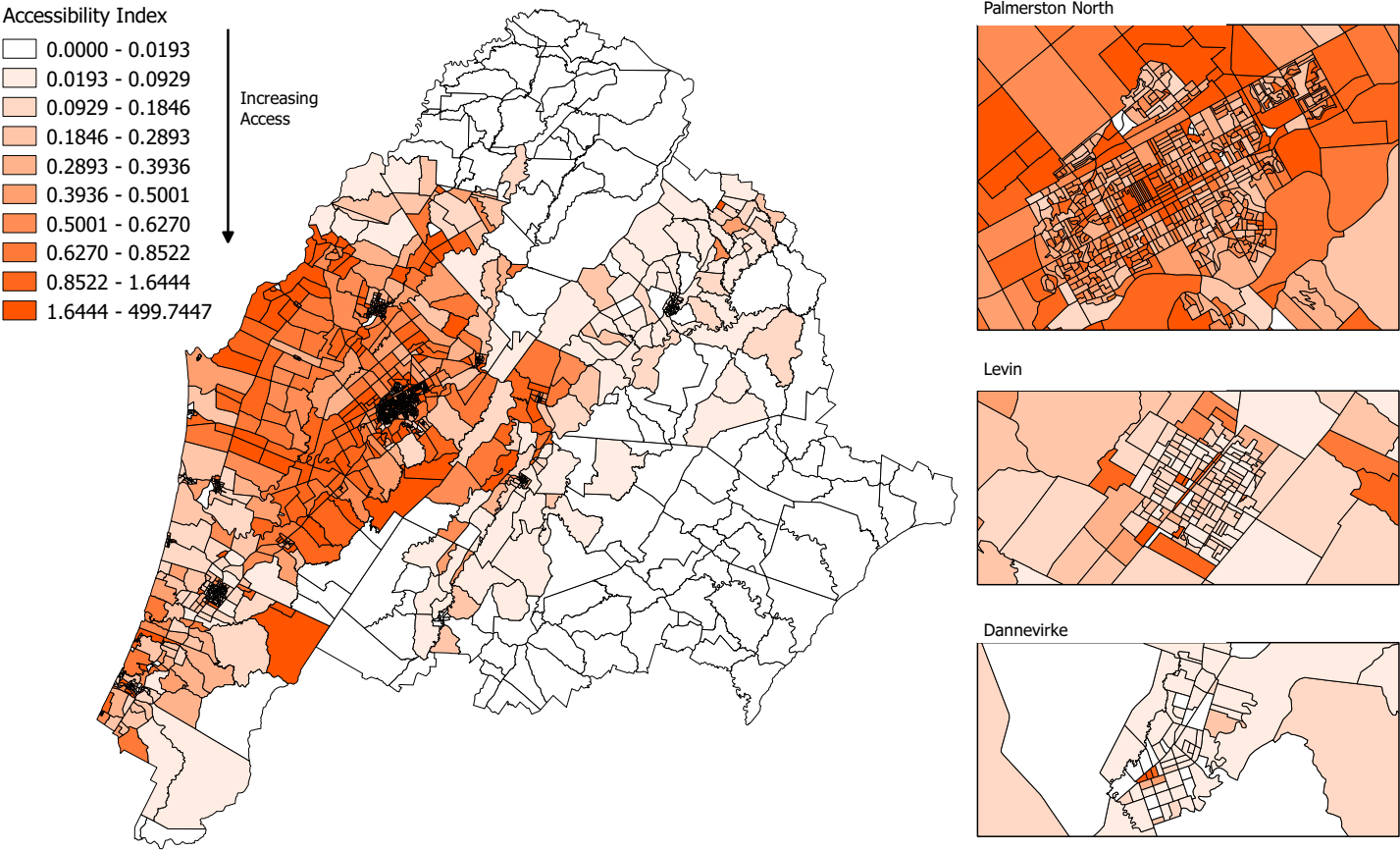


Figure 6.18: Accessibility for General Practices Using Two Step Floating Catchment With Distance Decay Method With Gorge Open

Figure 6.19 shows the percentage difference in the accessibility indices after the Manawatu Gorge was closed. The only area with a consistent pattern of decreased accessibility using this method was the area around Pahiatua.

The area directly surrounding Dannevirke showed a consistent decrease in accessibility pattern using the 2SFCA method. Using the 2SFCDD method the same area showed a mix of increasing and decreasing access, similar to the variation in the western areas of the district.

General Practitioner Two Step Floating Catchment with Distance Decay Ratio Differences Gorge Open versus Closed

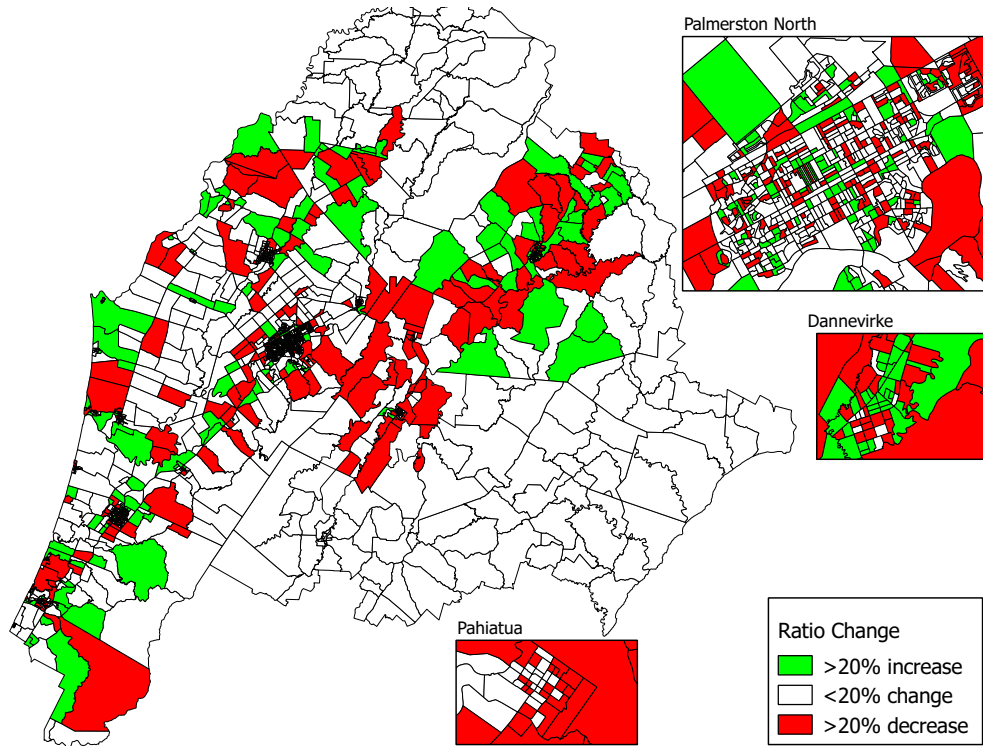


Figure 6.19: Accessibility Change for General Practice using Two Step Floating Catchment With Distance Decay After Gorge Closed

6.6 Realised Accessibility

6.6.1 Utilisation by Demographics

Figure 6.20 shows a matrix of utilisation rates by setting and demographics.

For all secondary services utilisation began to rise steadily after the age of 55. For GP utilisation the rise was more gradual from the age of about 30. Hospital outpatients exhibited a peak in utilisation rate 5-10 years younger than the other services.

The utilisation pattern by ethnicity for ED was distinct from the other settings. New Zealand European, NZ Maori and Pacific all exhibited roughly similar rates, while the Other ethnic group was much lower. This is distinct from the other settings where NZ European exhibited greater rates.

Utilisation by gender demonstrated that Females used GP services and were admitted to hospital at a higher rate than Males. Although there was a slightly higher rate for Females in Hospital Outpatient attendance it was only a small difference. For ED utilisation, Males exhibited a slightly higher rate than females.

All secondary services exhibited similar utilisation rate patterns by deprivation. Least deprived groups used these services at a lower rate than more deprived groups. In GP utilisation the deprivation 1 group used GP services much less than the other deprivation groups, while groups 2-4 exhibited much less difference.

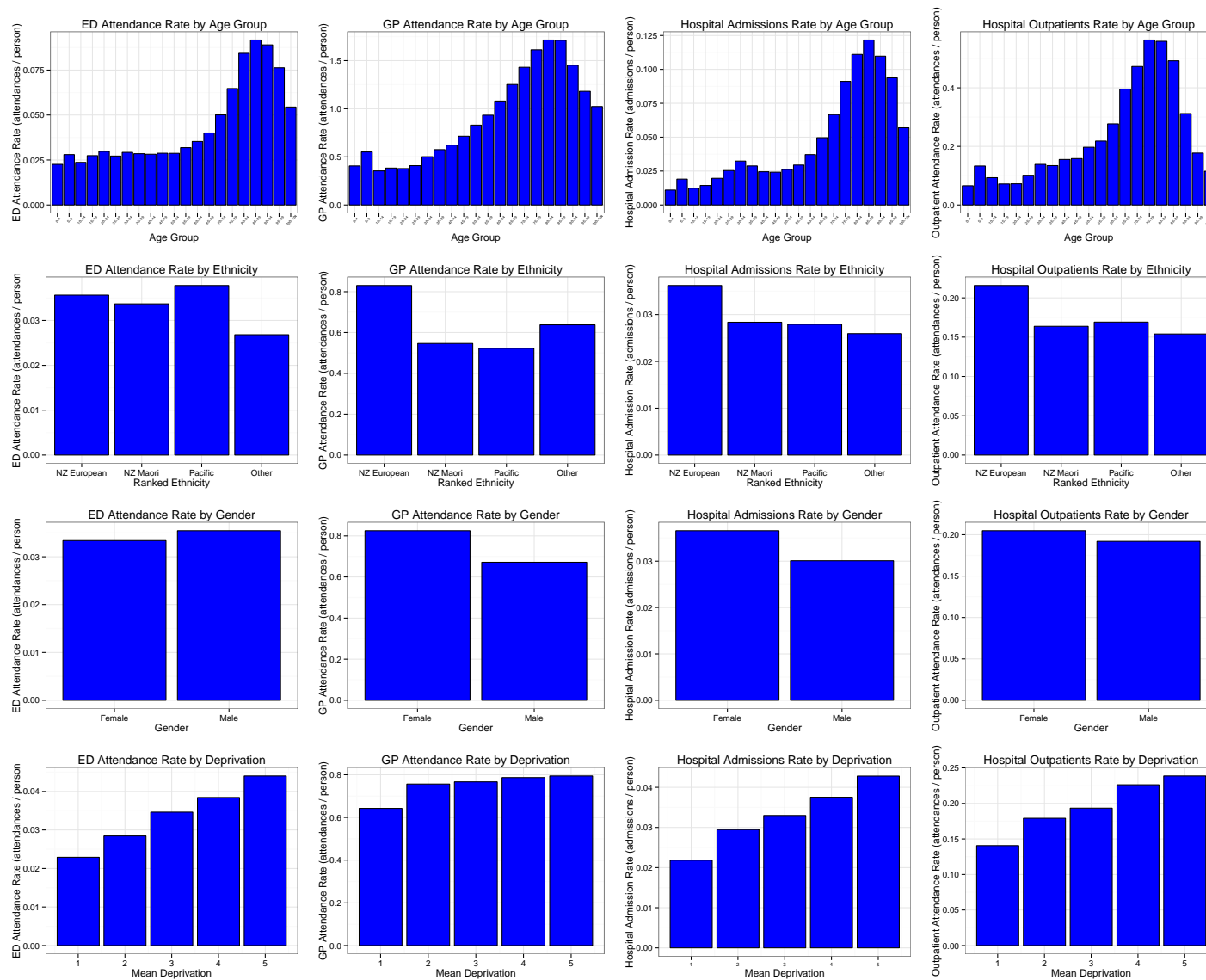


Figure 6.20: Demographics and Utilisation Rate Matrix

6.6.2 Choice

Figure 6.21 shows a kernel density plot of the number of practices in which patients were enrolled by meshblock. It represents realised patient choice. It shows the difference in choice exhibited between those in urban and rural areas. Those meshblocks that are rural had people that exhibited far less diversity in the practice they attend than those in urban areas. The vast majority of rural meshblocks had people who attended most commonly 1, 2 or 3 practices. Approximately 35% of rural meshblocks had populations of people that had no choice of service provider (Practice per Meshblock = 1). This is found in only 15% of urban meshblocks by comparison.

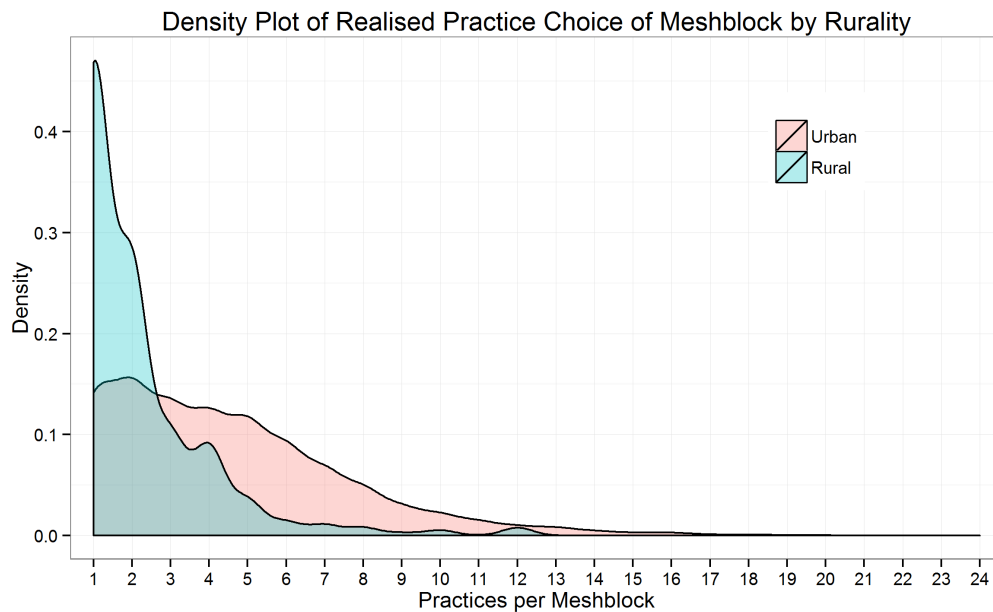


Figure 6.21: *Kernel Density of Practice Choice by Urban and Rural Meshblocks*

Figure 6.23 shows how this choice was distributed across the district. In rural areas choice was low, particularly in the east and north of the district. Choice remained relatively high in the west and southern areas. In the large urban centres of Palmerston North and Levin choice was high, while in Dannevirke, Otaki, Pahiatua and Woodville,

choice was low. Dannevirke, Otaki and Pahiatua all have large medical centres. Choice seems to be related to the number of medical centres in proximity to each meshblock, or expressed in another way, choice appears to be the realised measure of potential practice catchment.

Figure 6.22 shows how practice choice changed in each meshblock during the gorge closure period. There appears to be a large amount of variation in the western areas that should not have been affected by the gorge closure. In the eastern areas there was little change in practice choice.

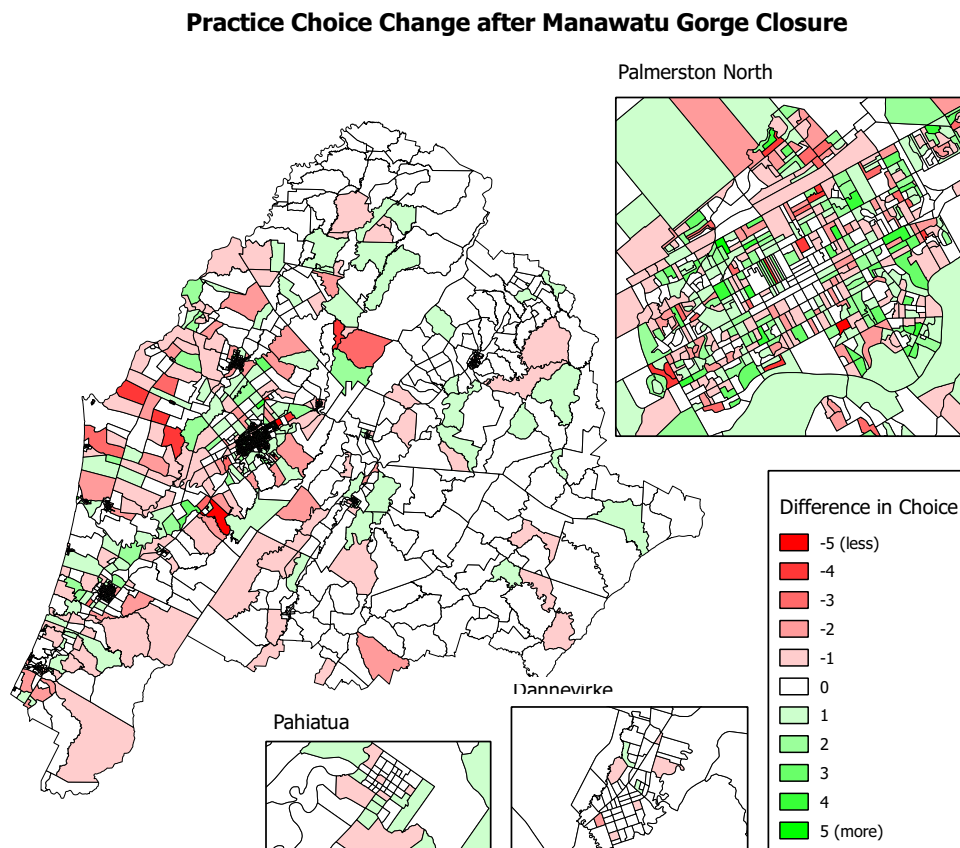


Figure 6.22: Accessibility Change for General Practice Using Practice Choice After Gorge Closed

There appears to be no association between choice and deprivation

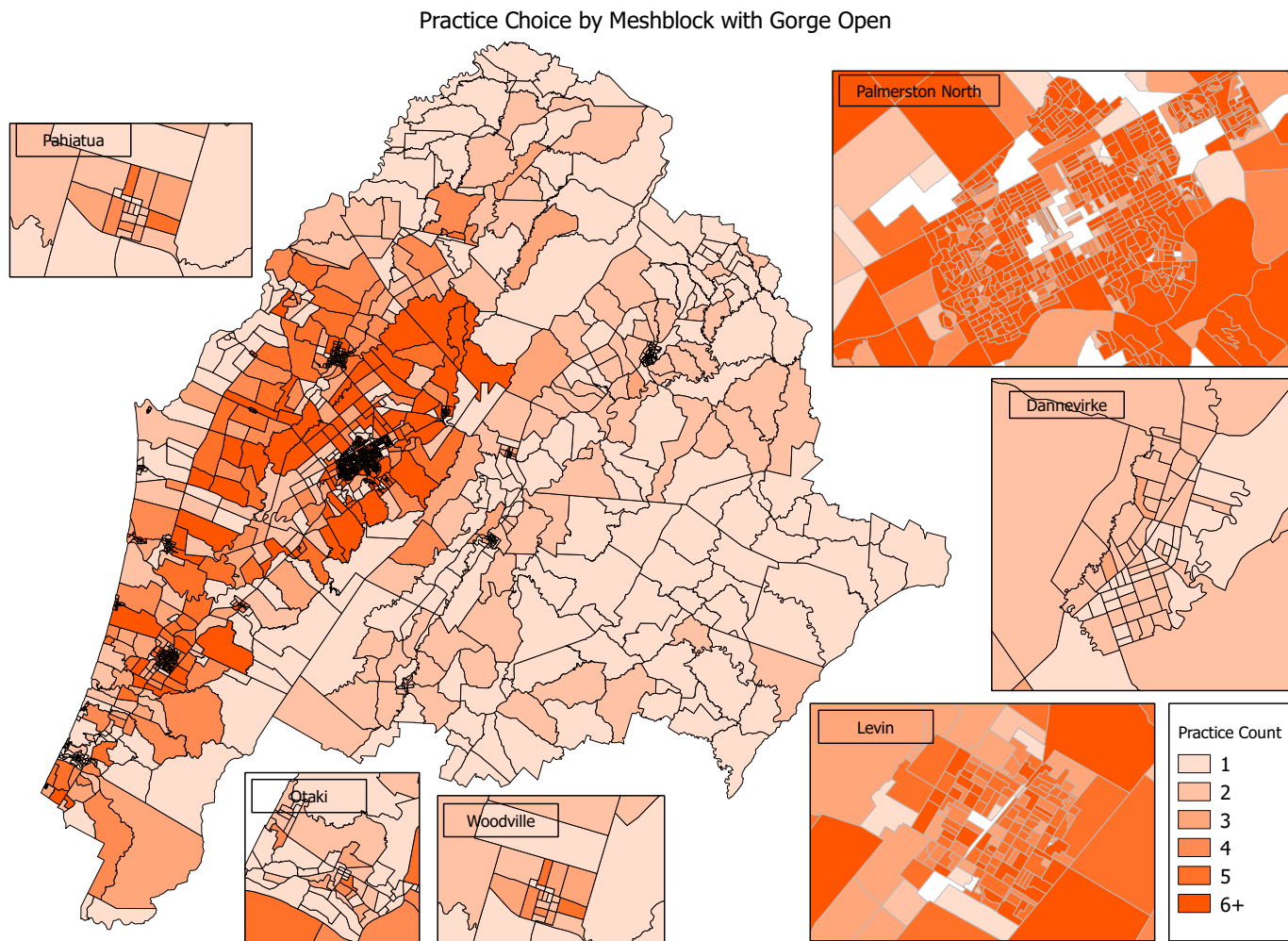


Figure 6.23: Accessibility for General Practices using Practice Choice with Gorge Open

(Figure 6.24). If plotted against maximum deprivation, there is a distinct relationship however (figure 6.25). For meshblocks with greater realised choice, there are lower maximum deprivation scores.

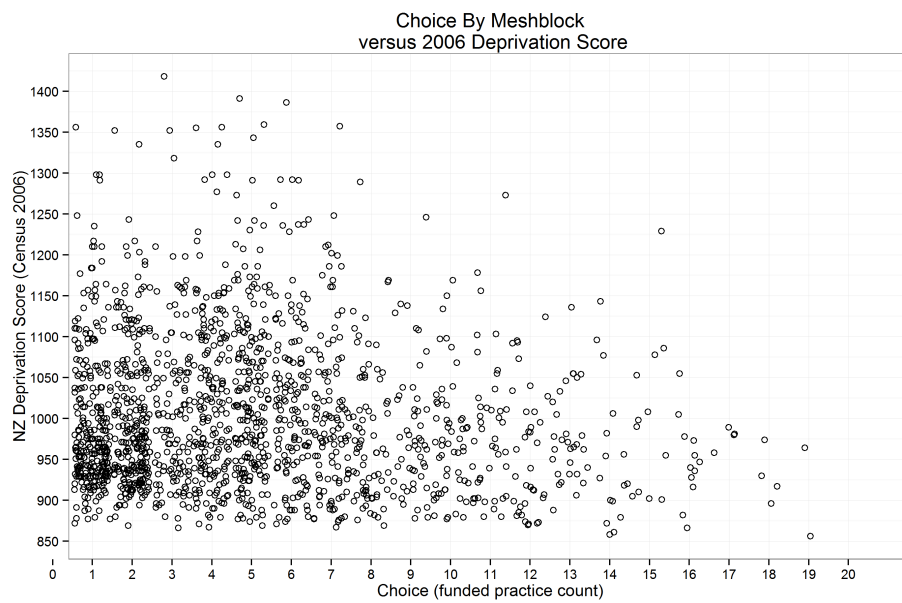


Figure 6.24: Practice Choice Compared With Meshblock Deprivation Score

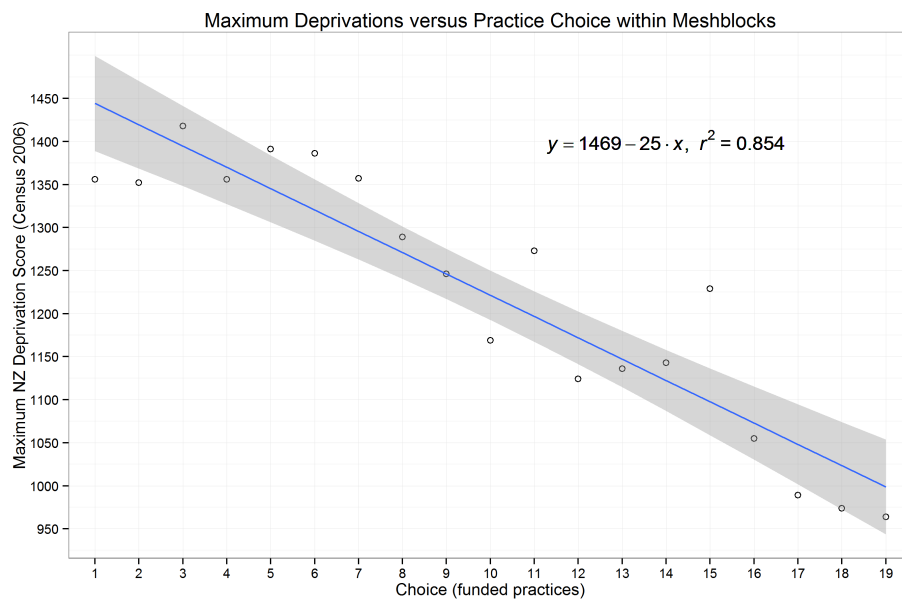


Figure 6.25: Practice Choice Regression With Maximum Meshblock Deprivation

6.6.3 *Intervention Analysis*

The results of a traditional travel-time category analysis can be seen in Figure 6.26. These show monthly utilisation rates for people living in each five-minute travel time category for each service with 95% confidence intervals. These results represent those from methodology typical in most contemporary literature that has investigated realised access to health services. They have no control or comparison data.

The results of the intervention analysis for both the control group of patients and the control period prior to the gorge closure can be seen for each service in Figure 6.27. Each graph represents the daily utilisation rate for each service. Green bars represent the period when the gorge was open while red bars represent the period the gorge was closed. The control group represents those patients who were not affected by a travel time increase to the service, while the intervention group were those patients that experienced an increase in travel time.

The rate-ratios for each of these services is summarised in Table 6.8 and Figure 6.28. The ratio is a convenient way to summarise the utilisation rates between the open and closed periods. A ratio below 1 indicates an increase in utilisation during the gorge closure and above 1 indicates a decrease. The p-value is the test statistic indicating a statistically significant increase or decrease in utilisation. Green bars in this instance represent the control group, while red bars represent the intervention group.

A summary of the results for each service is described in the following sections.

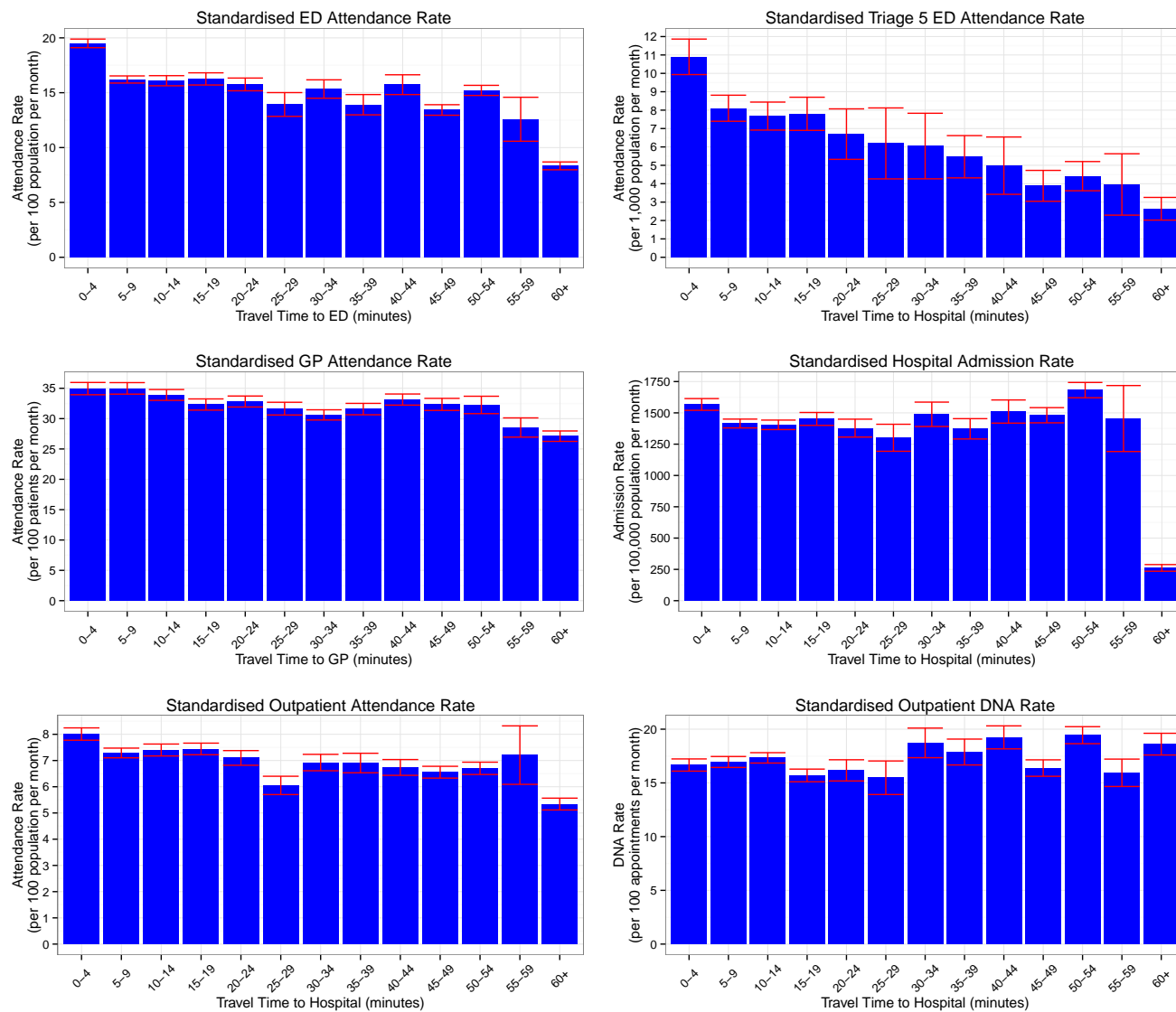


Figure 6.26: Travel Time Category Standardised Utilisation Rates with 95% Confidence Intervals

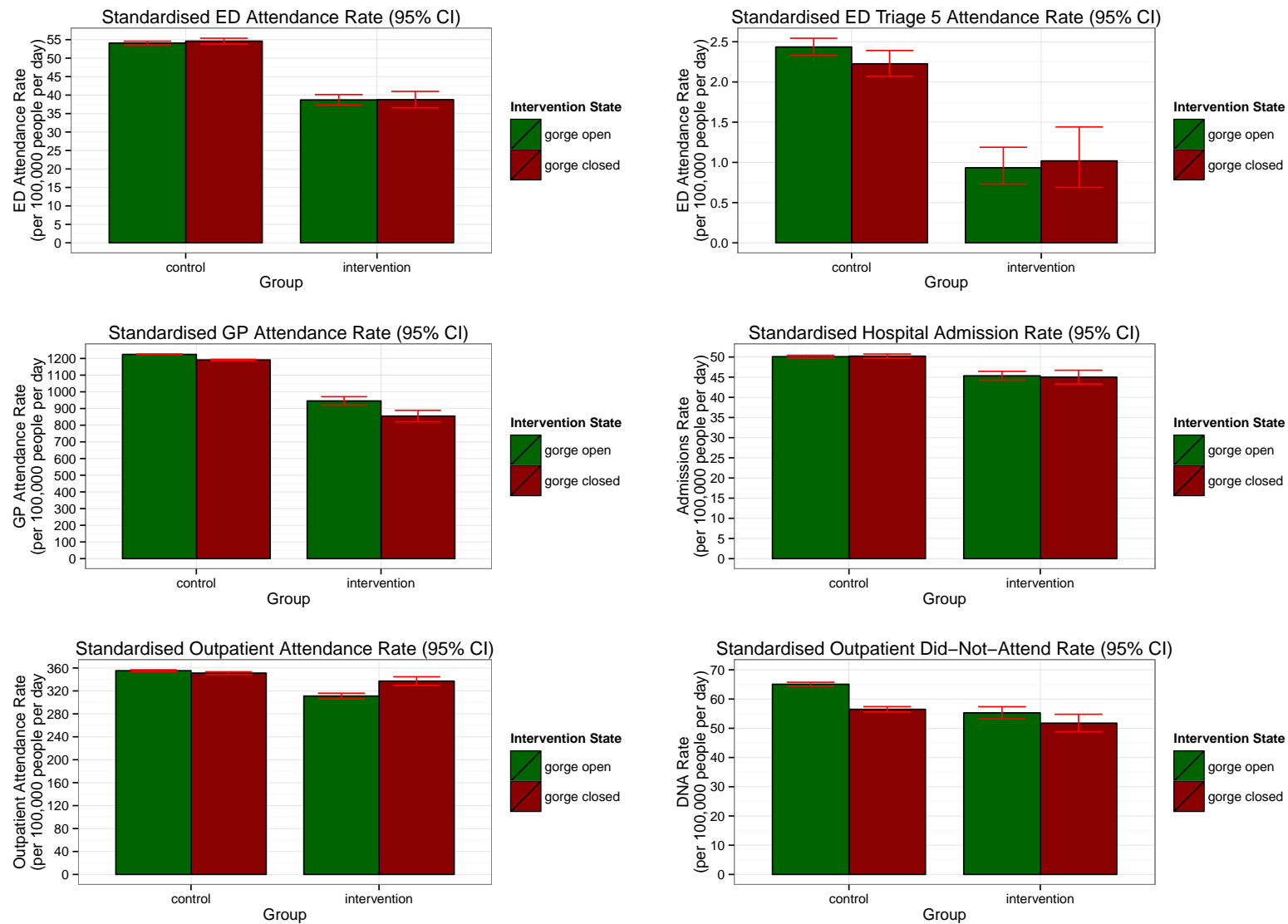


Figure 6.27: Standardised Utilisation Rate Comparisons for Control and Intervention Groups

Measure	Control				Intervention			
	ratio	lower ci	upper ci	p-value	ratio	lower ci	upper ci	p-value
ED Attendances	0.990	0.974	1.008	0.272	0.999	0.934	1.069	0.970
ED Attendances - Triage 5	1.093	1.006	1.189	0.035	0.922	0.609	1.429	0.711
Outpatient Attendances	1.010	1.004	1.02	0.004	0.922	0.898	0.9482	<0.001
Outpatient DNA	1.152	1.129	1.174	<0.001	1.068	0.998	1.1436	0.057
Hospital Admissions	0.998	0.985	1.01	0.702	1.008	0.964	1.05347	0.741
GP Attendances	1.028	1.024	1.031	<0.001	1.106	1.055	1.16	<0.001

Table 6.8: *Utilisation Rate Ratios with Gorge Open and Closed with 95% confidence intervals (shaded cells denote statistically significant results).*

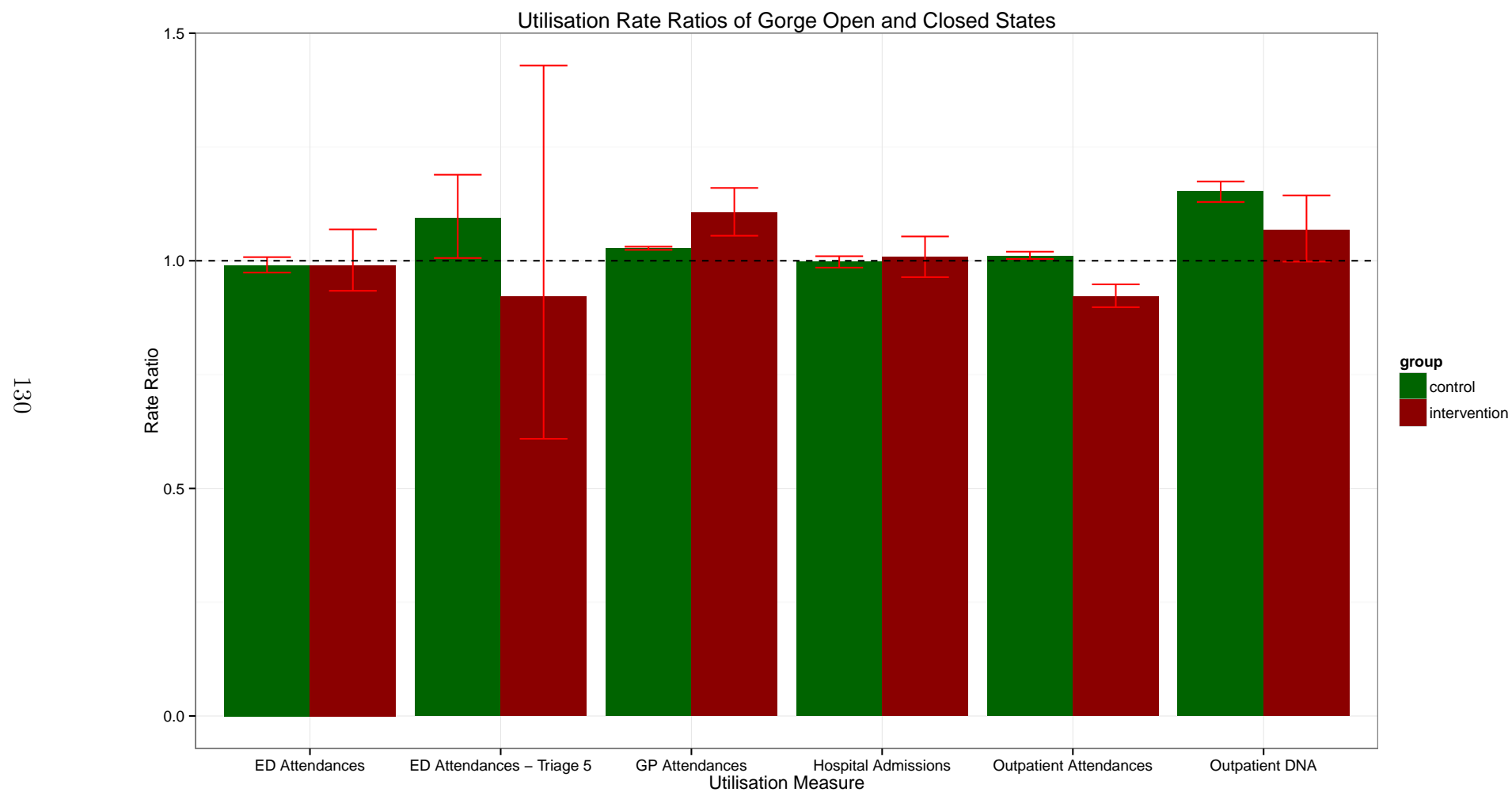


Figure 6.28: Utilisation rate ratios with 95% confidence intervals

6.6.4 ED Attendance

The attendance rate of the ED data from Palmerston North hospital can be seen in Table 6.9. There was a range between 24.2 and 25.4 attendances per 100 population per year between 2009 and 2012.

Year	Months	All Funded			Study Cohort		
		n	Visits	Annual Rate	n	Visits	Annual Rate
2009	6	150,675	16,774	0.22	101,456	9,385	0.19
2010	12	150,944	34,788	0.23	101,456	19,150	0.19
2011	12	152,287	34,582	0.23	101,456	19,469	0.19
2012	12	152,782	34,630	0.23	101,456	20,710	0.20

Table 6.9: ED Attendance Rates by Year for Funded and Cohort Populations

ED Attendance data conforms to a Poisson distribution. Figure 6.26 shows the attendance rate by travel-time categories with 95% confidence intervals. There appears to be four distinct bands of utilisation rate defined by travel-time boundaries.

Those less than five minutes drive to ED attend at the highest rate of all groups, with just under 19 attendances per 1,000 people per month. From five to 24 minutes there appears to be a step down in attendance rate to between 16 and 17 attendances per 1,000 people per month. This group is statistically distinct at a 95% level of confidence. The third band ranges from 25-59 minutes travel time. At this level attendance rates range from 13-15 per 1,000 people per month. The final band is for those who have a travel time of over 60 minutes to ED. At this level the attendance rate drops to just above 8, less than two times that of those living less than 5 minutes from ED.

This effect is even more pronounced when only those attendances that are the least urgent are analysed. Triage five attendances are those that are likely to be able to be treated in after hours or general practice. There appears to be a marked travel time effect, with those less than

five minutes from ED presenting at a statistically significant higher rate than any other travel time category. The decrease in rate continues for each travel time band. At sixty minutes the attendance rate for triage five attendances is four times less than at less than five minutes. Presentations that are triage five are the most likely to be presentations of convenience. The longer it takes to get to ED, the less likely it is that such a presentation would be convenient and this is perhaps reflected in the steady decline in access rates over with increasing time from ED. Those people 60 minutes from ED are almost certain to have a GP surgery less travel time away by virtue of the number of GP surgeries and their location through-out the district. If their condition does not require urgent attention it may then be more convenient for them to present to their general practice rather than face a length trip to ED.

When comparing utilisation rate differences between the control and intervention group while the gorge was open and closed there was no statistically significant difference. The rate ratio for each situation can be seen in Figure 6.28 with statistical analysis shown in Table 6.8. Both rate ratios were below 1 indicating a higher rate of attendance during the gorge closure, however their confidence intervals cross 1 and p-values were above the 0.05 level. Consequently I accept the null hypothesis in this test that there is no difference in access rates of control and intervention groups before and during the gorge closure period for ED attendances.

ED attendance for those assessed as being triage level 5 (least urgent) showed a decrease in access in the control group and no significant change in the intervention group. The change in the control group was statistically significant at the 5% level.

6.6.5 Outpatient Attendance

Table 6.10 shows the appointment rate for outpatient appointments was 1.1 per person in both the funded and cohort populations. The only exception to this rate was in 2010 in the funded population where

the rate was 1.2 appointments per person.

Year	Months	All Funded			Study Cohort		
		n	Visits	Annual Rate	n	Visits	Annual Rate
2009	6	150,675	84,676	1.1	101,456	55,123	1.1
2010	12	150,944	174,550	1.2	101,456	114,615	1.1
2011	12	152,287	171,641	1.1	101,456	114,624	1.1
2012	12	152,782	162,533	1.1	101,456	115,215	1.1

Table 6.10: Outpatient Appointment Crude Rates by Year for Funded and Cohort Populations

Outpatient attendance and DNA data conforms to a Poisson distribution.

Analysis of outpatient attendance rates by five minutes travel time categories (Figure 6.26 on page 127) reveals a pattern of approximately three access rate bands. Those that live less than five minutes travel time to the hospital present at around eight attendances per 100 population. Between five and 59 minutes there is a rate of approximately seven presentations per 100 and those over 60 minutes have a rate of six per 100. The over 60 minute group have a statistically significant lower travel time than all other groups, while the less than five minute group are statistically significant from all other groups except the 55-59 minute group.

Outpatient attendance access decreased for the control group while increasing for the intervention group (Figure 6.28 and Table 6.8). Both results were statistically significant at the 5% level.

Outpatient DNA rates across travel time categories were lower in the 0-29 minute groups at about 16 DNAs per appointment. After 30 minutes there is a trend of higher DNAs peaking at approximately 19 DNAs per appointment.

Outpatient DNAs decreased for the control group, while remaining unchanged for the intervention group. The p-value of the intervention

group change (0.057) is very close to the 0.05 threshold however.

6.6.6 Hospital Admissions

Table 6.11 shows the crude hospital admission rate seen in the data set used for this study for all funded patients and for those in the study cohort. The rate for the cohort was lower than for the funded population and showed an 1% increase each year from 2009, while the funded population admission rate remained unchanged.

Year	Months	All Funded			Study Cohort		
		n	Visits	Annual Rate	n	Visits	Annual Rate
2009	6	150,675	16,414	0.22	101,456	8,740	0.17
2010	12	150,944	33,233	0.22	101,456	18,071	0.18
2011	12	152,287	33,882	0.22	101,456	19,405	0.19
2012	12	152,782	32,363	0.21	101,456	20,695	0.20

Table 6.11: Hospital Admission Raw Rates by Year for Funded and Cohort Populations

Hospital admissions data conformed to a Poisson distribution.

The admission rate by travel time category (Figure 6.26 on page 127) shows those living less than five minutes from the hospital have a higher admission rate than those living up to 29 minutes away where it remains relatively constant. Between 30 minutes and 59 minutes there appears to be an increase in admission rates. Those living 60 minutes or more away exhibit a very substantial drop in admission rate, almost five times less than the closest group of patients.

There was no change in hospital admission rates in the control or intervention groups.

The case weighted discharge values exhibit a log-normal distribution. Table 6.12 shows a comparison of the means of CWD for each group and the p-values for the comparison of log-normal means between each

Gorge State Group	Open	Closed	p-value
Control	0.850	0.866	<0.01
Intervention	0.878	0.883	0.29
p-value	0.01	0.06	

Table 6.12: Case-Weighted Discharge Means for Control and Intervention Groups with Gorge Open and Closed and t-test results of log-normal distributions (shaded cells indicate statistically significant results).

group. Case weighted discharge values increased in the control group while the Gorge was closed, while no subsequent change was seen in the intervention group. There was a difference between the CWD values between the control and intervention group while the gorge was open with the intervention group exhibiting higher CWD values, but not while closed.

The control group experienced an increase in case weighted discharge value while the gorge was closed. There was no subsequent significant increase in the intervention group.

6.6.7 GP Attendance

The crude rate of GP attendance ranged from 4.3 to 4.4 visits annually per patient across all funded patients, while it ranged from 4.4 to 4.6 visits annual for the study cohort (Table 6.13). Both rates followed a very slight downward trend over three and a half years.

GP attendance data conforms to a Poisson distribution.

Figure 6.26 (page 127) shows that those living less than 15 minutes from their GP service exhibit the highest use by travel time category. Those living between 15 and 54 minutes appear to exhibit a similar rate. People living more than 55 minutes have the lowest rate.

GP Attendances decreased in the control and intervention groups (Fig-

Year	Months	All Funded			Study Cohort		
		n	Visits	Annual Rate	n	Visits	Annual Rate
2009	6	150,675	332,705	4.4	101,456	232,161	4.6
2010	12	150,944	661,578	4.4	101,456	453,487	4.5
2011	12	152,287	665,960	4.4	101,456	450,255	4.4
2012	12	152,782	660,656	4.3	101,456	452,128	4.5

Table 6.13: GP Attendances Over Study Period for All Funded and Cohort Patients

ure 6.27 on page 128) and both were statistically significant (Table 6.8 on page 129) . The decrease was greater in the intervention group, with the confidence intervals of the control and intervention group providing mutually exclusive ranges.

Chapter VII

Discussion

One of the strengths of this thesis is that it uses a natural experiment to reduce selection bias that is often seen in prior studies. Astell-Burt et al. (2011) comment on two prior studies (Monnet et al., 2006, 2008) that looked at the geographic aspects of patients seeking hepatitis C treatment. They are skeptical of these studies' findings because of the lack of any validated controls for deprivation in light of findings in their own study that controlling for deprivation has a significant effect on decreasing the apparent impact of travel-time. They suggest that in conditions where there are social stigmas associated with a particular condition and in small communities such as those seen rurally, there may be a barrier from fear of being exposed or perceptions of reduced confidentiality within the community. They suggest this acts as a confounding factor when trying to measure the effects of travel times to typically urban treatment centres. Fuller et al. (2000) identifies a culture of self-reliance as a contributing factor to differences in health seeking behaviour seen in rural versus urban settings. Hoyt et al. (1997) have demonstrated that the effect of stigmatism occurs in those that live in rural towns or villages, but is less likely for those that live outside such settlements on farms and outlying properties.

When measuring use of health services by distance it is difficult to compare populations. Groups that form travel time categories or controls are often composed of individuals of varying rural status. This means that even with all other factors considered (age, ethnicity, deprivation, gender) most measures will not dissociate the effect of rurality versus

travel time to services.

Rural meshblocks are larger than urban meshblocks (Table 2.1 on page 27). This is because of the way meshblocks are composed to contain roughly equal populations. The different geometries of meshblocks can mean that generalisations applied to them by using population level data can create confounding conditions.

This thesis attempts to control for the inherent differences between those who live close to and far from health services. For calculations of potential accessibility individual patient point data is used to calculate travel times. Geographical generalisations are not used even when data is being aggregated to meshblock or larger units. Realised accessibility calculations have been made between a control and intervention group across two periods of time. The purpose of assessing the period of time prior to the gorge closure and while it is closed is to look at the same population of people and how their use of health services changed. Having the control group gives an insight into the effect factors extraneous to travel time have had upon the whole population. This is a unique approach among existing literature on the effect of travel time on health care accessibility.

7.1 *Free and Open Source Software and The Scientific Method*

The term 'Free and Open Source Software' (FOSS) encapsulates two related concepts (Fuggetta, 2003). The term 'free' refers to the liberty of anyone to do anything that they wish to with the software. These liberties may include the ability to run the software; to study how it works; to change it; and to re-distribute it in either a changed or unchanged state. 'Free' in this context does not imply cost. There is much software that is cost-free but contains restrictions on how it may be used. This software is often referred to as 'freeware'. The term 'open source' refers to the state of the underlying source-code of the software being available to anyone to inspect it. It is possible for software to be open-source but to carry with it restrictions on how it

may be used. Free software implies it is open source but the converse does not apply. It also implies that there is no cost associated with the intellectual property of the software. The term 'Free and Open Source' appears pleonastic but it serves the purpose to differentiate between that software which is simply free of cost with that which imposes no restrictions to liberties on its use and it differentiates it from that software which is open source but continues to have use restrictions.

The scientific method and FOSS are complimentary. FOSS and collective scientific knowledge are developed through the same 'bazaar' principle described by Raymond (1998). He describes the 'bazaar' as:

...a great babbling bazaar of differing agendas and approaches ..
out of which a coherent and stable system could seemingly emerge
only by a succession of miracles

Scientists publish their findings in academic journals to share their results publicly with a wider community. These works are often the adaptation of others' prior work. FOSS is built on the principle of sharing source code and having the liberty to freely adapt this to new needs. Science requires that results are reproducible. FOSS does not restrict scientists creatively or financially.

Geographic information systems have traditionally been commercial proprietary products that are expensive to license. A large body of existing GIS literature uses ARC GIS to perform network analysis, to calculate travel times across road networks and to visualise geospatial data (Martin et al., 1998; Brabyn and Skelly, 2001; Bagheri et al., 2005; Lovett et al., 2002; Parker and Campbell, 1998; Messina et al., 2006). Commercial and proprietary products present a financial barrier for researchers to recreate research and findings of others.

This thesis has used FOSS exclusively for GIS calculations (PostgreSQL, PostGIS and pgRouting) and visualisations (Quantum GIS) and statistical analysis and presentation (R). By doing so I have demonstrated that these products can be used in a health setting with large data sets

to make accurate and meaningful analysis. This has implications for the health sector, particularly in New Zealand where a large proportion of non-governmental agencies providing services have charitable status and may not have the funds to invest in expensive commercial or proprietary products. Even calculations on large data sets can be made using modest desktop computing hardware with these tools.

The methods and algorithms described here can also easily be recreated by using the same FOSS platform by other researchers and institutions.

7.2 Geocoding

7.2.1 Last-Best-Known Geocode

The geocoding process applied by the Ministry of Health to PHO register addresses appeared to be less than optimal. It didn't appear to deal with relatively minor imperfections in address data. Given the same inputs, the quality of the information it produced appeared to vary over time.

The Last-Best-Known Geocode process (LBKGP) is computationally simple. It will run on modest hardware with standard relational database products over large data sets. Relational databases are designed to perform uniqueness and matching operations over large sets of data efficiently. This results in a fast performing process to instantly improve the data quality of any data set presently geocoded with the MOH PHO register process.

The improvement is such that anyone working with health data sets and using PHO registers as the basis for geospatial data should undertake the described LBKGP to improve their data quality and sample sizes by as much as 37%.

7.2.2 Corrective Geocoding

Despite the improvement in geocoding by using LBKGP there were still a substantial number of addresses that remained uncoded. There were 15,110 addresses representing 20,939 patients that still did not have a geocode with sufficient certainty to assign meshblock information. This represents 10.8% and 11.2% of total addresses and patients respectively from all Central PHO registers from 1 July 2009 to 1 October 2012.

Validation Standards

The Corrective Geocoding Algorithm (CGA) performed variably dependent on the matching method used.

The CGA's ability to correct addresses was validated against manual human calculations. Although this is a time consuming exercise it is possible to complete for a modest number of records.

The CGA's ability to provide accurate geocode coordinates was compared with the results from the Google geocoding application programming interface (API). The gold standard for validating the CGA geocoding would have been to manually determine the spatial location of each address used in the validation set by taking a GPS reading for each address' corrected physical location. Traveling to each address would have been a time and cost prohibitive exercise even over a small validation sample (220 records).

Google provides a free geocoding service via a well documented API. Access to this service is free but Google places limits on its use, including a transactional speed and volume limit (Google, 2013). These limits are relatively modest, allowing only 2,500 uses of the API per day for free access, as well as limiting use for displaying on a Google map. This study had 28,505 addresses that needed geocoding. Using the Google API to geocode the study data was not a viable option. It was however practical to use it for the much smaller 220 record vali-

dation set. Using the Google interface in lieu of the gold standard is justified on these grounds.

The Google API would not have been a viable candidate for correcting addresses as a replacement for the described CGA. Aside from the volume limits Google imposes on the use of its API, it is unclear how well the Google API corrects spelling and typographical errors if at all. It was therefore important to ensure the addresses that were passed to the API were correct and accurate. There are potential privacy issues with supplying address information of entire patient registers to a service such as Google. Google is a business built on advertising revenue. It is in their interest to store queries to search and geocoding systems for use in delivering targeted advertising to users. As such, supplying complete address data on entire patient registers would not be inadvisable.

Performance

The largest number of matches resulted from method 5 (see Table 6.2 on page 88) which used a phonetic match to road name and exact matches to road type and suburb or city name. This suggests that at least 21% of addresses that do not geocode in the MOH process are as a result of minor spelling errors in road names. The positive predictive value of method 5 was high at 96%, and the mean location error for correctly matched roads was only 465 metres. The results that method 5 match wrongly however had a substantial error.

The next largest matching method was method 1 resolving 16% of records with a PPV of 100 and a mean location error of 94 metres. Methods 2 and 4 also provided PPVs of 100 but contributed much smaller proportions of matches.

A few hundred meters in terms of travel time is only 30 seconds even at low motor vehicle speeds. Such an error is tolerable if the patient sample sizes can be increased significantly. For the purposes of determining meshblock and corresponding deprivation quintile such an error

is small and would be unlikely to affect this study in any appreciable way. Methods 1,3 and 4 differed by less than 100 metres on average compared with the Google results. Although method 5 differed by 465 metres this was a sizeable difference for the travel time calculations I attempted to make. Method 2 exhibited an high variance in location for correctly coded addresses from the location reported by Google (6.1 km) and was excluded from the final correction process on these grounds.

When the algorithm incorrectly codes addresses, it does so with substantial error. The overall mean error for all incorrectly classified addresses was 66.4 kilometres. Method 3 is the first method where addresses are classified incorrectly, and although the accuracy for correct addresses in method 3 is high at 92 metre, for those that are incorrect the mean error is 23.5 kilometres. At high speeds, 23.5 kilometres represents approximately 15 minutes of travel time. This error margin was high given that the travel time increases caused by the gorge closure were 10-15 minutes. Method 5 contributes similar error when addresses are incorrectly matched, although its PPV is greater than method 3.

Based on these considerations I chose to use methods 1,4 and 5 of the CGA to correct uncoded addresses. Weighing up the total number of addresses each method corrected and their respective PPVs and locational errors these methods represented the best balance of volume of addresses corrected to precision of those corrections.

Overall using this corrective geocoding technique was successful in allowing a large number of additional patients to be eligible for inclusion in the study, that may otherwise not have been.

Summary

Using both the LBKGP and the CGA I coded 57.7% of the original 21,902 addresses that were not geocoded in the standard MOH process. Those patients eligible for inclusion in the study increased by 20.5% by applying both the LBKGP and CGA. Using the LBKG I coded an

additional 6,794 addresses. In total 7.0% of the 140,250 total unique addresses in the study period remain uncoded. It is likely a substantial proportion of address location information was also improved by the LBKGP by correcting issues similar to those illustrated in Table 4.2 on (page 50).

Applying a corrective geocoding process to the data is a relatively novel approach in the literature. Fortney et al. (1999) used GPS data from patient physical addresses to improve the data quality of geocoding but this was for a very small study size of 106 patients half of whom already had addresses with valid geocodes. Such an approach cannot be applied to a large data set easily.

A more common approach in the literature to account for poor address information is to map less specific address data to postal codes. Point data is calculated by using geographic centroids of each postcode area (Athas et al., 2000). For rural areas in New Zealand this approach can introduce substantial error into travel time calculations.

The methods that I have developed and described as part of this study can be applied to large data sets while still maintaining the precision of point calculations.

Further experimentation of the parameters used to correct addresses may yield even better results than seen here. Different combinations of the various matching components and parameters such as the distance tolerance could be analysed to find a more optimal matching method. Applying the match methods in a different order, or independently could yield different results. Further work should be completed to refine this approach.

7.3 Travel Time Models

Travel time or distance can be calculated either using Euclidean or Road Network methods. Euclidean calculations are more simple to implement. There is mixed evidence about how well Euclidean distance

correlates to actual driving distance or travel times (Phibbs and Luft, 1995; Jordan et al., 2004; Haynes et al., 2006), but some authors speculate that the ability to use Euclidean distance accurately as a proxy for network travel distance depends on local geography (Witten et al., 2003). Because Euclidean distances use only source and destination points to make their calculations, any change in physical barriers between the start and end points of a journey cannot be reflected in their measurements. The Manawatu Gorge has changed the travel distance and time for people to health services, but Euclidean calculations cannot represent these. For this reason, road network travel time analysis is the only appropriate approach to take in this study.

Brabyn and Skelly (2001) provides the best travel time model in the literature for estimating travel times from patient domiciles to health services in New Zealand. This approach has been explicitly used in many publications (Brabyn and Skelly, 2001, 2002; Brabyn, 2002; Brabyn and Barnett, 2004; Bagheri et al., 2005, 2008; Beere and Brabyn, 2006), while others have not provided exact implementation details but suggest similar approaches (Pearce et al., 2007). It uses the New Zealand Topographical 50:1 (NZTopo50) map data to provide road network information. This data set has limited attributes to establish accurate road speed estimates however.

The validation of the Brabyn-Skelly model by Lauder et al. (2001) suggests it may be a good predictor of travel time on New Zealand roads. Care must be taken in interpreting these results. The validation against New Zealand Automobile Association (AA) data included trips up to and over 20 hours long, used arc lengths for sinuosity calculations different to that described in Brabyn and Skelly (2001), used a different threshold for straight versus bendy roads and GPS data was limited to a geographical area in New Zealand's most populated region. From the graphs provided it appears that only a small proportion of observations were for trips of less than 60 minutes. Most of the population of New Zealand have hospital services within 45 minutes drive of their domicile (Brabyn and Skelly, 2001). Validating the travel time model on a

majority of trips that exceed this threshold does not express how well the model performs for the majority of calculations that will be made using it. Automobile Association travel times include stops for each hour of travel. Most people traveling directly to access health services would be unlikely to stop because of the length of trip being usually less than 60 minutes. This would also suggest that the Brabyn-Skelly model's 6.6% slower travel times compared to AA estimates are likely conservative. The authors acknowledge that the arc length calculations made during the validation are longer than those described by (Brabyn and Skelly, 2001) and that they used a different threshold (1.046 versus 1.02) to determine straight versus bendy roads. Increasing arc length is likely to reduce the relative sinuosity, moving more roads into the 'straight' category. Increasing the sinuosity threshold will also move more roads into the 'straight' category. Both of these are likely to have had the effect of providing faster travel speeds than Brabyn and Skelly's 2001 original implementation. Overall the results of the validation from Lauder et al. (2001) suggests the Brabyn-Skelly travel time model will provide slower travel speeds than observed times.

Lauder et al. (2001) provides interesting information on the variation in measured velocity of travel across roads of different sinuosities. Variation increases markedly with increasing sinuosity up to an index of 1.06. This effect may be exaggerated by long arc lengths used in sinuosity calculations and may also reflect the varying skill and confidence levels of drivers through windy roads. The implication of this information is that travel time variation across highly sinuous routes may be so high as to make it difficult to accurately model travel times without understanding and accounting for the other factors that affect these variations.

Some travel time models use parameters to predict travel speeds based on whether they are in an urban or rural environment. There are numerous taxonomies used to define whether environments are rural (Hewitt, 1989). Although speed limits in urban areas will tend to be lower, there are additional constraints that may result in people trav-

eling slower than the posted speed limit. These include intersections, traffic lights and traffic controls such as stop and give way signs. Urban environments also have a higher number of distractions and hazards such as footpaths, driveways and signage. All of these factors can cause drivers to travel more slowly as they take in and process the busy environment.

Deriving attributes in travel time calculations is intended to provide a travel time model with parameters which act as proxies for the physical environment or conditions that may influence driving speeds but are not present in the existing attribute data.

Brabyn and Skelly (2001) derived road sinuosity from calculations on the road network data, and urban boundaries by using ground coverage and applied these to the NZTopo50 data. Very little other literature that classifies roads uses derived attributes. In order to get the most detailed road network information possible I combined the NZTopo50 and NZOGPS data into one network data set. From this I derived a number of additional attributes to those that Brabyn and Skelly (2001) used.

7.3.1 Intersection Density

My intent in deriving intersection density was to use this as a potential replacement for the urban parameter Brabyn and Skelly (2001) derived from ground cover or the urban areas as defined by Statistics NZ.

Although Brabyn and Skelly (2001) cited the reference to the method they used to derive urban boundaries from ground cover, I was not able to source the description of the process. I have therefore not recreated that as part of this study.

The bounds of the Statistics NZ urban areas (see Figure 6.1 on page 90) appear to be large and encompasses road networks that would not have the typical properties of an urban environment. These urban areas appears to extend into farmland and lifestyle blocks. This boundary also

suggests that there is an urban corridor between Palmerston North and Ashhurst which does not exist. Statistics New Zealand urban area information is not useful for the purpose of defining urban road networks.

I examined different intersection density levels to establish which visually approximated the high density road networks associated with an urban environment. I used 500 metre grids as the area unit to match the journey segment lengths used in other derived attribute calculations. I was looking for the density at which the pattern of grids was such that there were very few grids in rural areas and most urban grids were contiguous.

There were a large number of isolated grids in rural areas that had densities of one or two. At a density of four, there were several gaps in central Palmerston North and Feilding. A density of three appeared to approximate my expectation of urban area the best. This was therefore the threshold I used. This equates to a density of twelve intersections per square kilometer.

The area with a high intersection density is much smaller than the Statistics NZ urban boundaries (Figure 6.1 on page 90). The area between Ashhurst and Palmerston North has few areas of high intersection density. Feilding, Woodville and Pahiatua all have small areas of high intersection density around their urban centres.

At the threshold I chose there were a number of grids that exhibited high intersection density in relatively rural and remote areas. These are caused mainly by artifacts of how the road network is expressed as geometries.

Deriving intersection density for travel time calculations appears to be a novel approach, particularly in health care applications. It provides results that appear sufficiently accurate to use as a proxy for inclusion in a travel time model.

7.3.2 *Elevation*

There is a marked difference in the elevation profiles of the Manawatu gorge compared with the two alternative routes of the Saddle Road and Pahiatua Track. Both alternative routes traverse mountain ranges, while the gorge road makes use of the natural water gap of the Manawatu River.

Most travel time models are two dimensional. Three dimensional models incorporate a z-coordinate which indicates elevation. There are no freely available road network data sets in New Zealand that include z-coordinate information. Traditional travel time models have not factored in elevation into their calculations.

I wanted to include elevation data into a travel time model to determine whether it had any significant effect on travel times, especially given the substantial difference in elevation of the routes of Manawatu Gorge, Saddle Road and Pahiatua Track. Elevation may have an effect on driving characteristics. Vehicles traveling uphill may be inclined to accelerate more slowly and have a lower maximum speed due to the extra load that is placed on a vehicle's engine in moving the weight of the vehicle and passengers uphill. Alternatively maximum speeds up hills may be increased because of gravity-assisted shorter braking distances. Vehicles traveling downhill may have an increased braking distance as brakes must counter-act both the vehicle's kinetic energy and the released potential energy of its descent. Increased braking distances are likely to result in slower speeds into and through corners. Acceleration is likely increased on declines.

Predicted elevation results using the nearest neighbour and interpolation algorithms were not satisfactory (Figure 6.2 on page 91). The Manawatu Gorge is likely a difficult use case for predicting road elevation from contour data. The road is proximal to the side of a steep gorge with a considerable drop away to the Manawatu river below. Contour lines within the gorge are very close together which provides much less margin for variation in either the contour or road data set

and related calculations. Proximity to such dramatic changes in elevation are likely what is causing some of the inconsistent results from interpolating from contour data. Applying the same approach to roads in more open environments may be more appropriate, but perhaps less useful as in less extreme environments road elevation is unlikely to be such a considerable factor in travel time calculations.

The purpose of developing a travel time model is to ensure it is generalisable to other areas not directly observed or measured. Persistence in using the 20 metre contour model with more sophisticated interpolation or other techniques may have yielded better results to predict elevation. Such detailed investigation of one small component of the wider study was not justified, particularly in light that there remains no evidence that including elevation in a travel time model improves it. No other travel time model I have found in the literature uses road elevation or gradient as a component of travel time calculations. I therefore did not include this attribute in the final model calculation.

7.3.3 Maximum and Total Turning Angle

The sinuosity index of a road is a gross determination for describing how a road deviates from a straight line. Sinuosity indices describe nothing about the road except the overall ratio between the straight line distance and the distance following the road path. They are a composite measure to represent the sharpness of turns and the number of turns in a road together. It is not possible to differentiate whether a road has a large number of gentle bends, or a small number of sharp bends from a sinuosity index.

Figure 7.1 shows two different shapes. Shape A has a sinuosity index of 1.4 and shape B has an index of 1.5. These indices are relatively similar and would indicate roads that are both bendy. It is possible however to describe these shapes using alternative measures that reflect their differences.

The shapes can be described using the angles formed by their composite

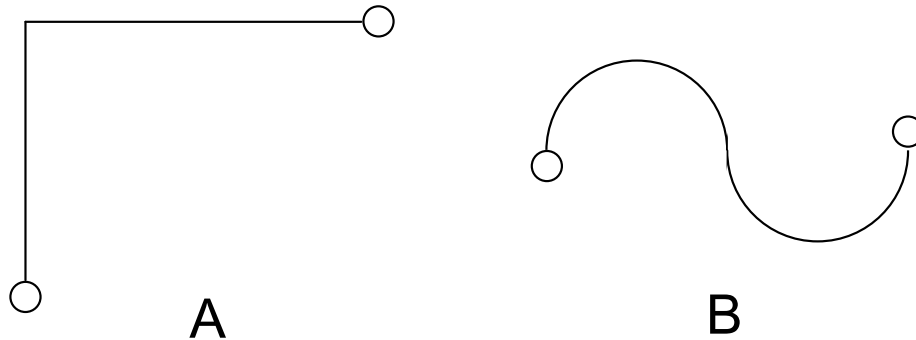


Figure 7.1: *Example of differing road shapes*

lines. There are two properties which describe the angular nature of each shape uniquely. The maximum angle represents that portion of the line which has the sharpest angle. The total angle represents the sum of all angles the lines form.

Shape A has a maximum angle of 90 degrees, while shape B has a maximum angle of about 12 degrees. This second calculation is highly dependent upon how the curve is composed of separate straight line segments. In this case each 180 degrees of curvature has been broken into a component of 15 straight-line segments similar to the way in which road network data is composed.

Shape A has a total turning angle of 90 degrees, while shape B has a total turning angle of 360 degrees. It is possible to have total turning angles higher than 360 degrees for particularly windy roads. Because total turning angles are a sum of all angles within a segment of road they will increase proportionately to the arc lengths used in road calculations in much the same way that sinuosity is affected by arc length.

Both maximum and total turning angles can describe properties of roads that the sinuosity index is unable to. Using these discrete measures it is possible to classify roads with similar sinuosities into very distinct categories. The approach of using maximum and total angles appears to be novel in the literature.

The angle at which a motor vehicle must be turned through is important in determining its maximum speed. The speed at which cars negotiate roads at differing angles depends on factors inherent to the vehicle, the camber of the road, the road surface and the inherent skill of the driver. Roads with a higher maximum turning angle will be slower to negotiate because of the need to decelerate into and accelerate out of each turn. I included both maximum and total turning angles as a derived attribute in the travel time model in an attempt to account for such factors beyond a basic sinuosity index.

7.3.4 Linear Model

After examining the plots of speed for each road attribute (Figure 6.4 on page 95) I thought that a multiple variable linear model may be able to predict travel times. There appeared to be some linear relationships particularly with maximum speeds and attribute values.

Maximum Sinuosity index and intersection density both exhibit linear relationships. Total turning angle appears to exhibit two linear relationships for values below 500 degrees, and for those above 500 degrees.

I tried various transformations for each attribute's data to give a more linear association with speed. Intersection density was the only attribute that transformed to a more linear state. In this case it appeared to fit a log-linear model. I therefore chose to use the logarithm of intersection density in the multiple linear regression calculations.

The AIC stepwise regression process ranks and excludes least significant variables. Those variables that remained in the final model included Speed Class, Road Class, Sinuosity Index and the logarithm of Intersection Density. All dummy variables were excluded. Although Lanes were not excluded, they were not significant at the 5% level in the model.

The derived model appeared intuitive. Sinuosity Index and Intersection Density decreased the predicted speed as their values increase. Sinuos-

ity Index has the single largest influence on speed of all variables. As Speed and Road classes increase they increase the predicted speed. All attribute affects were applied to the base model speed of 50.2 km/hr (the intercept in the regression analysis). This itself appeared intuitive given it is the predominant urban area speed limit.

The model had an adjusted r^2 value of only 0.6667. Although this indicates a weak correlation to the data my initial testing of the model against known routes produced wildly extreme results for travel time predictions.

The potential problem with this type of linear model is that it cannot represent the complexity of the effect attributes such as sinuosity have on travel at different speeds. The nature of the model is that the effect of sinuosity is applied as a constant to the travel time prediction. This however doesn't model how traffic is affected by sinuosity in reality. At lower speeds, sinuosity will have less of an impact on travel speed than at higher speeds because there will be fewer bends where a vehicle must slow down to turn a corner.

Attempts at predicting travel time using linear models may be more successful by building models specific to each speed class. This would alleviate the problem of attributes being applied constantly across all speed classes. It would provide an opportunity for sinuosity to affect high speed roads more than lower speed roads. I didn't attempt to create such a model as this concept only became clearer to me towards the end of the study on reflecting on my findings.

7.3.5 Rule-Based Travel Time

After the linear regression model was unsuccessful I approached the problem of developing a new travel time model from a different perspective. The assumption that I made was that only the most constraining attribute has an impact on travel speed. By examining the relationship between each attribute value and speed for each speed class I was able

to derive rules that constrained speed for each attribute. I also determined that a number of attributes appeared to constrain the minimum speed also.

This approach is effectively an extension of that used by other travel time models that assign road segments into classes which define their speeds (Brabyn and Skelly, 2001; Christie and Fone, 2003; Haynes et al., 2006). I developed rules which gave each road an upper and lower constraint with the resultant speed determined by taking a mean of those values. This resulted in 188 different speed profiles applied to roads in this travel time model.

I applied an adjustment factor to results based on the speed class of the road for which calculations were being made. The factor applied was always below zero which decreased the final predicted speed. The purpose of this factor was to bring model maximum predicted speeds back within the legal speed limit. The model was developed based on GPS observation data which explained the occasional higher than speed limit predictions.

Early testing of this model indicated that it had promise in predicting road travel times by car.

7.3.6 Model Comparisons

Model performance can be compared relative to each other. Such a comparison can give an insight into the characteristics, strengths and weaknesses of each model. This can be done at a large scale with little effort, as empirical data does not need to be collected as the comparisons are simply between other models. I undertook the model comparison as an initial step to test how well the MacRae model performed against both the Brabyn-Skelly model and Google's travel time predictions.

I needed to implement the Brabyn-Skelly model to compare it with the model I developed. While I tried to reproduce as closely as possible the

Brabyn-Skelly algorithm described there were some minor differences in the implementation from that published. Brabyn and Skelly (2001) created road segments (they referred to them as arcs) that were approximately 500 metres in length by joining individual road segments together but they were unclear how they dealt with segments that were a composite of two different road classes (e.g. speed, road type, rurality etc). My implementation uses the slowest common class when a segment is composed of multiple classes. Brabyn and Skelly (2001) used ground coverage to determine urban areas. I substituted ground coverage with Statistics NZ urban boundary data. I did this by spatially intersecting road lines with the urban areas. Any road segment that intersected an urban area was marked with an urban flag.

Both of these differences in my implementation are likely to cause the model to predict slower travel speeds. It is unclear what the magnitude of this effect would be in relation to mixed attribute journey segments because Brabyn and Skelly (2001) didn't describe their method for this particular aspect of their model. Because the Statistics NZ urban areas are very large and appear to extend well beyond urban development it is likely that roads in my implementation will be flagged as rural that may not have been in the original implementation. This would certainly have a net effect of predicting longer travel times.

Where one model has been validated it can be used as a reference standard which can reduce the need to measure new models against empirical data. Although the Brabyn-Skelly model has been validated, the validation methodology had some potential flaws (Section 7.3). I chose to use two additional models as references because of these potential problems and the lack of any other validated models. The additional reference models were a basic travel time calculation (called the Basic model here) and Google's travel time predictions. The Basic model was based on road speed limits which provided the theoretical minimum legal travel time. Google travel time predictions were determined using the Google web API.

All models were highly correlated to one another (Figure 6.4 on page

95). Model correlation analysis is a gross generalisation and doesn't reflect the characteristics of each model and how they perform at different distances. All models exhibited a correlation that was slower than the Basic model which was expected. The Basic model represents the theoretical minimum travel time between points without traveling faster than the speed limit. In practice journeys will be slower than this because of other traffic traveling slowly, stop signs and traffic lights and the need to slow below speed limits to turn corners.

Lauder et al. (2001) performed correlations on travel time data but forced the linear model to have a zero intercept. The premise of this was to calibrate the models at the starting point and reduce anomalous situations where low travel times between compared models have large proportionate differences due to non-zero intercept values. I chose not to do this, as the purpose of the correlation is not to predict travel times from one model to another, but to understand the relationships between the models themselves. It is therefore not necessary in this comparison to ensure travel times between models are rational compared to each other.

The correlation coefficient between the Basic and Google predictions was exactly one with an r^2 of 0.973. This indicates a strong correlation where the predictions of travel time that Google made increased proportionately to the predictions of the Basic model. This may suggest that the Google travel time calculations are weighted using speed limits as a significant component of their prediction model.

The correlation coefficients of the Google compared to the MacRae model suggest that overall the MacRae model is likely to be slower. The correlation has a very high r^2 so it is reasonable to predict from a coefficient of 1.2 that the MacRae model will be 20% slower than the Google model overall. A negative intercept suggests that the MacRae model will be faster over shorter distances. A similar pattern exists between the Brabyn-Skelly and Google models, but with the Brabyn-Skelly model being approximately 40% slower.

The performance of the models at the 30 and 60 minute thresholds are important for assessing access to health services. These are the thresholds used in catchment calculations to measure performance of health services against contracts. Thirty and sixty minutes of travel using the Brabyn-Skelly model will equate to 24 and 50 minutes using the MacRae model. Catchments calculated with the MacRae travel time model will be geographically larger and therefore include a greater population.

These observations are confirmed when predicted travel times from all models are plotted on the same graph (Figure 6.5 on page 96). The closest approximation to the Basic travel time is seen in the Google model. This is consistent with the results of the model regression analysis. The relationships between the models is not strictly linear. The Brabyn-Skelly models make similar predictions for trips of between 100-120 minutes.

This insight into the model differences suggests that when compared with empirical data, there will be a difference in how the Brabyn-Skelly and MacRae models perform for trips under 100 minutes and over 120 minutes.

7.3.7 Model Validation

The intent of validating the travel time model was to understand how well the model performed against real-world data compared to other models. Model comparisons suggested that each model had differences in how they predicted travel times.

The method that I employed to collect observation data had some limitations. I chose not to include unsealed roads because of the low volume of traffic that would travel along such roads. This would have required a significant investment of time to collect sufficient data to make the observations meaningful for the purpose of validation. I was unable to reliably collect travel time information in dense urban areas. The method I used to collect observation data involved determining vehicle

travel times between two points. Vehicles in a dense urban situation have many choices in the routes they can take between two points. Observations of travel times across two points in an urban area would have a large loss to follow-up. Observing more than two points would have required more resource than I had for this exercise. Rather than use data which may have contained significant bias I opted to exclude it from the validation. Observation data relied on identifying the same vehicle across two separate videos which restricted observations to daylight hours.

The variance between predictions and observations was small for short trips compared with long trips. Not validating the model of shorter urban trips therefore likely has had little implication for this study.

The issue of not validating the model over unsealed roads does have potential implications for this study. A proportion of those that live at the periphery of catchments or in rural areas are likely to use some unsealed roads to travel to health services. Most travel to general practice and hospital outpatient services and planned admissions are done during the day. Presentation to emergency departments could be done at any time, but some attendances at ED come via ambulance services rather than private motor vehicle. Not validating the models in night-time driving conditions has likely had little impact in assessing spatial barriers in this study.

How night time driving affects travel time is unclear. Driving speeds may be slower due to increased difficulty in seeing the road and hazards. To offset this the nature of other impediments such as traffic volumes and traffic lights, and associated decreased waiting times at intersections may make journeys faster. Ultimately care needs to be taken in interpreting the validation results for use outside daylight conditions.

The travel times observed for each site and direction of travel appear to have different variances (Figure 6.6 on page 98). The sites with the least variance were the journeys between Otaki and Waikanae and over the Haywards Hill road. The largest variances were seen in the journeys

between Palmerston North and Ashhurst and over the Saddle Track.

The variance in observed travel times at different sites cannot be easily explained by either trip length or sinuosity of the roads. The Palmerston North to Ashhurst route is straight open road similar to the Otaki to Waikanae route. Yet they both sit at extremes of variance of observed travel times. It is possible that the variance seen is more a function of opportunities for traffic to travel at the speed at which they are comfortable. Such opportunities may be passing lanes or lower traffic densities that either allow the freedom of travelers to move at their own pace, or greater opportunities to overtake due to less oncoming traffic. Where fewer opportunities are present more traffic may have to travel to speeds lower than their maximum ability. This has the consequence of compressing variation in observed travel times. This result contrasts to observations that have previously suggested more sinuous roads have higher variation in travel speeds (Lauder et al., 2001).

Median travel times for each site also varied depending on the direction of travel. The reason that this may happen could also be related to the same reasons set out for the variance in travel times across sites generally. The direction of travel on the road network may influence the number and nature of opportunities for motorists to travel at the maximum speed they are comfortable.

The performance of the Brabyn-Skelly model is consistent with previous validation of this method (Lauder et al., 2001). It over-estimated journey times compared to observations. This is consistent despite some methodological differences in how the algorithm was implemented. This suggests that the differences in implementation had little effect on the predictions the model made.

The MacRae model performed generally well but performed the worst of all models for predicting time traveling over the Rimutaka Hill road. The Rimutaka Hill road is very sinuous. This road did not form part of the road network sample during the collection of data for the development of the travel time model. The extreme nature of the road com-

bined with it not being factored into the development of the MacRae travel time model may be a contributing factor to why the MacRae model performed poorly compared with other models in predicting travel time across it.

The Ashhurst and Hayward Hill roads were not included in the data for development of the MacRae model either. The model predicted these travel times with considerable accuracy however. These roads are far less extreme in terms of their characteristics and it is likely roads similar to these were included in the model development data. The Saddle Road was included in the model and it also under-estimated the travel time.

The poor prediction of the Rimutaka Hill Road times may suggest that the MacRae travel time model is not generalisable. Its performance against routes that were part of the data collection that formed the basis of the model development may be due to a training bias effect. It has performed well against observations that were not included in sample data collection for model development also. This suggests that a more conservative view of these results could be that caution should be exercised when applying the model to roads with extreme attributes. The MacRae model has a tendency to predict shorter than observed travel times across roads that are more sinuous.

When representing the overall error of the models by minutes (Figure 6.7 on page 99) the Brabyn-Skelly performed the worst. The median error however being only 1.3 minutes suggests reasonable accuracy. This metric in isolation would suggest the model could be appropriate for our purposes. The biggest issue with the model is that the nature of the way in which it overestimates journeys is such that the median percentage error is high at 15.9%. The Brabyn-Skelly model had more error in short journeys.

The MacRae model performed very well having a median error in minutes of -0.5 minutes and an inter-quartile range of 1.6 minutes. It had the least inter-quartile range of all models in both gross and percentage

error.

Predictions made by Google were marginally the most accurate with the MacRae model being a close second. The large variation in both gross and percentage error in the Google model show this is the least precise.

Although travel-times as predicted by Google were generally the most accurate, they performed the worst for the Saddle road observations. At the time the traffic survey data was being collected from the Saddle Road, the New Zealand Transport Agency had already undertaken a number of road upgrades to this section of road due to its increased use. The road network data that I used for calculation of both Brabyn-Skelly and MacRae models was updated weekly, and likely factored in the improved driving conditions through this area. It is possible the information on which Google based its travel time calculations had not been updated with such information resulting in the slower predicted journey. Using the Google predicted travel times would not have been feasible given the licensing restrictions placed on the use of the Google data. Regardless of licensing restrictions the model may not have been the best choice for estimating travel times . The fact that the Saddle Road was the main bypass for the Manawatu Gorge and the majority of altered journeys during the gorge closure would have gone through this route could have introduced substantial error into travel time calculations.

Most people round their travel time estimations to intervals of five to ten minutes (Haynes et al., 2006). The upper quartile of gross error of all models was under 5 minutes. It is likely given this that the errors seen in the models fall within the range of rounding error applied in human perception and reporting of travel times.

The MacRae model performed particularly well over the Pahiatua Track and reasonably over the Saddle Road. Both of these are the alternative routes for the Manawatu Gorge used when it was closed. I decided to use the MacRae travel-time model for the basis of my calculations

for this thesis given the evidence that it is generalisable for all but the most sinuous roads and given its better accuracy and precision compared with the Brabyn-Skelly model.

7.4 Cohort Demographics and Selection

Age is the most variable demographic when comparing the profiles of the funded population to those included in the study cohort (Fig 6.8 on page 101).

The inclusion criteria requiring continuous funding during the entire study period is likely to contribute to a substantial portion of the differences seen between the funded population and selected cohort. The study period begins on 1 July 2009 and ends on 31 December 2012, totaling a period of three and a half years. Funding is calculated each three months. Patient registers are submitted on or around the 20th of the month prior to the funding period start to allow time for processing and payments to be calculated. This means that the pertinent cut-off dates for those included in the study are 20 June 2009 and 20 August 2012.

Anyone born after 20 June 2009 (the cut-off date for the first funding period of 1 July 2009) would not be funded for the whole of the study period and would therefore be excluded. This would account for the large difference in the proportion of those in the 0-5 age group. This age group has effectively been reduced from a five year band to one year and six months because of this inclusion criteria. Children would need to have been born before 20 June 2009 but after 1 January 2009 to be included in this age group and the study. This is approximately 30% the size of other age bands.

Anyone moving to the district after the 20 June 2009, or leaving the district prior to 20 August 2012 would not be funded for the whole of the study period. Palmerston North is the major centre for Massey University, a tertiary institution catering to a population of students

predominantly 18-39 years old (Ministry of Education, 2013). The majority of undergraduate degrees in New Zealand are three years long. Universities are only situated in six urban areas in New Zealand and many institutions offer specialised courses. This contributes to a high proportion of university students migrating temporarily to these centres. It is likely that this pattern contributes to the lower proportion of those between the ages of 20-25 seen in the cohort. As students complete their tertiary studies, usually between the ages of 21-23 they will move in search of employment. For a large proportion of those in this situation, that would involve moving out of Palmerston North and the MidCentral District. A number of students would be lost to follow-up resulting in this age range having a disproportionately high number of exclusions.

Patients must register with a general practice every three years to remain funded. This is a relatively trivial task for patients that visit the general practice with a frequency more than this. When a patient presents and they are close to their three year window, there are usually processes in each general practice to re-enroll patients. A large number of practices have processes established to identify patients who are close to having their funding cut because of this rule, and these practices will contact them by post, have them complete an enrollment form. This postal enrollment is less likely to be successful in groups of people that have more transient domiciles and who do not attend their general practice frequently. This could explain the high numbers of people in this age group who have been excluded because those in the 25-35 age group are both likely to have a high proportion of people that attend general practice infrequently because they are healthy and they are more likely to live in flatting situations and therefore have more transient domiciles. The opposite is likely to be impacting the 50-55 age group. This group of people are likely to visit the general practitioner more frequently and have a more settled domicile.

Those people over the age of 89 are likely excluded from the study at a higher rate than the general population due to the substantial increase

in mortality rates with increasing age. Anyone that died between 20 June 2009 and 20 August 2012 would have been excluded from the study because they were not funded continuously for the study period.

The effect that this inclusion criteria had on the cohort age profile was negated by age adjusting utilisation rates. The inclusion criteria that required continuous funding was important to ensure that when comparing control and intervention group rates, both groups consisted of the same individuals for the period of the study.

Ethnicity of the cohort was closely representative of the NZ population. There was a slightly higher proportion of NZ European and slightly lower Pacific and Other compared to the 2006 Census. The lower Pacific proportion compared with the national average is not surprising given the density of Pacific peoples living in the Auckland urban area, which would raise the national average. The rate at which NZ Maori and Other ethnicities were excluded was higher proportionately, while NZ European was lower. The exact reason for this effect is unclear, although it may be related more to the age profiles of the different ethnic groups than anything else (Figure 6.8). There is a higher proportion of the NZ Maori population in the 20-35 age groups. This is the age group that is excluded at a higher rate. Because of this those with an ethnicity of NZ Maori will be excluded more than NZ Europeans by virtue of their having proportionately more people in that particular demographic.

NZ European and Other ethnicities live predominantly in rural areas in MidCentral. Given that ethnicity is correlated with both where people live (Figure 6.11 and Figure 6.12) and how much they use health services (figure 6.20) there is a need to control for these potentially confounding factors. NZ Maori and NZ Pacific use health services except ED at a greater rate than other ethnic groups. They dwell in urban areas at a higher rate than other ethnicities and therefore by proportion will live closer to health services in general. All of these factors can impact an analysis of utilisation of health services by travel time.

Demographic characteristics are not uniformly distributed across geographic areas. People live in communities which can self-select for particularly demographic attributes. The geographic distribution of demographic profiles has relevance in this study because the way in which the control and intervention groups are selected is effectively by geography. The area affected by the Manawatu Gorge closure is limited to a single geographic area (Figure 6.14).

Vast swathes of Palmerston North, Dannevirke, Levin, Otaki and Pahiatua have the highest deprivation profiles. Deprivation is calculated on the census population of each meshblock using a number of proxy measures, which include access to a motor vehicle and access to a telephone. Access to both of these can impact how people seek health care. Without a motor vehicle people must either rely on telephone consultations or use public transport. For those that live in urban areas, access to a motor vehicle is not as essential for day-to-day living as those that live rurally. The same logic applies for telephone access, where sheer driving distance means that not having a telephone would represent a higher hardship for rural living.

The least deprived areas appear to be on the outskirts of urban centres. This may be caused by more affluent urban dwellers that can afford more land to move to the outskirts of urban areas onto lifestyle blocks. In Palmerston North, these areas are still relatively close to Palmerston North hospital and urban general practices.

New Zealand Maori and Pacific have almost identical deprivation and age profiles as do NZ European and Other. It may be possible that age and deprivation are associated in some way, contributing to these similarities.

7.5 *Potential Accessibility*

Potential accessibility to health services can describe the opportunities that a population has to use those services. It requires only access to

population data and provider information. This data is less sensitive than patient health data and in most circumstances more available. Population counts can often be sourced from census type data. Facility counts can usually be determined from publicly available information and provider count data has little privacy implications.

There are a variety of ways to measure spatial access to health services. Each method has advantages and disadvantages, and these must be considered in the context of what is being measured and what assumptions are being made (Higgs, 2004).

The purpose of this thesis is to understand the impact the closure of the Manawatu Gorge has had on the accessibility of health services. Potential accessibility measures therefore need to be considered in the context of the accessibility they present while the gorge is open and closed to both hospital and general practice services.

7.5.1 Travel Time Isochrones

Travel Time isochrones show the travel time boundaries from a geographical point (see Figure 6.14 on page 109). Isochrones are limited in that they can represent travel time from only one point.

The five minute travel time categories I used were arbitrary. There is evidence that patients will tend to round travel times to five or ten minute intervals (Haynes et al., 2006). Having isochrones in such categories may match patient perception of travel times to services.

I chose to calculate isochrones up to sixty minutes only. This represents the prescribed after hours travel time catchment.

The isochrones appear to provide a detailed view of travel time. Corridors of extended travel distances represent faster road speeds. These can clearly be seen running along state highways. This suggests that the MacRae model used to generate travel times has sufficient sensitivity to detect the effect a road such as a state highway can have on

travel times. It also illustrates how isochrones can illustrate particularly localized changes in travel time boundaries.

The impact of slower urban travel speeds can be seen in Palmerston North, with the western and south-western aspects of the city further from the hospital by time than some semi-rural northern areas (Figure 6.15 on page 110). This is likely because the hospital is located in the north-eastern quadrant of the city. Again this can be useful in understanding the accessibility on a relatively small scale such as an urban centre.

The area affected by the gorge closure is predominantly the area to the east of the Manawatu Gorge. This area represents that in which travel times for residents moved into a higher category while the gorge was closed. Both Dannevirke and Woodville were affected. Residents of these towns would normally use the Gorge as the fastest route to Palmerston North. Pahiatua sits on the fringe of the area affected. Residents south of Pahiatua would normally find it faster to travel to Palmerston North via the Pahiatua track rather than going through the Manawatu Gorge. This is because the location of the Gorge is north of the track and the additional time taken to travel to the Gorge offsets the savings of the faster road network.

Travel time isochrones do not take into account service resource. They calculate travel times from one service point. They are a relatively intuitive measure, and show the distribution of populations and how much time it takes them to reach services by road. They were useful in defining the geographical area potentially affected by the gorge closure. They are too simplistic a measure to use to analyse multiple service points such as general practice.

7.5.2 Provider-to-Population Ratio

The provider-to-population (PTP) ratio method is also known as a supply ratio (Guagliardo, 2004) or demand-supply ratio (Luo and Wang,

2003). It uses a ratio of the supply of services compared to the population within defined geographic boundaries. This method is applied typically to large geographical units because it does not account well for border crossing of patient populations. It has also been used on smaller geographical areas, combined with more advanced techniques to provide service density maps (Teach et al., 2006).

A Provider-to-Population Ratio can either be expressed as a ratio with providers or patients as the denominator. It can be a relatively intuitive and straight-forward figure for people to understand. The greatest downfall of this method is its reliance on counts of service units and its inability to cope with arbitrary geopolitical boundaries and how patients cross these (Love and Lindquist, 1995).

District Health Boards have clearly defined geographical boundaries and therefore a PTP ratio can be used to measure district wide health care accessibility.

District Health Board boundaries in New Zealand are often located close to large urban populations. For example Wellington and Lower Hutt are two cities whose city centres are located only 16km apart yet have separate DHBs. There is substantial potential for patients to seek health services across DHB boundaries when large urban populations are close to each other. The term for this event is 'inter-district flow'. In New Zealand there are no geographical limitations on where a patient may seek their care.

Supply ratios are problematic when you have populations of people that cross geographic boundaries to seek health services however. In traditional PTP methods, patients are only counted in one geography, usually based on their domicile. When a patient uses services in more than one geography, they are using supply somewhere where it is not being accounted for in the equation.

The inter-district flow of patients must be considered in this study. This issue is of particular concern in Otaki which lies at the southern boarder of the district and only 10 minutes drive from the Kapiti Coast towns

of Waikanae and Paraparaumu, both of which are in a different DHB catchment to MidCentral's. Using PTP ratios to measure accessibility in this situation may result in a skewing of figures.

When the movement of patients from one geographical area to another is unbalanced, the area with the net gain of patients underestimates demand and the area with the net loss overestimates it. It is unclear how this relationship plays out in Otaki and therefore what affect this has on the PTP ratio. Because the ratio is calculated across the whole district it is likely that this effect would not be substantial. There may be larger implications if sub-district boundaries were used for PTP ratios.

Two measurements are required to calculate a PTP ratio; a representation of services available; and a representation of the population. Health services can be represented by different units. The largest may be health facilities, such as hospitals or general practices. The smallest may be expressed in provider-appointments or bed-days. The most appropriate unit to use will depend on the situation being investigated.

Large units are easier to count. There can be substantial disproportion between the resources facilities have. Using small units such as provider-appointments or bed-days requires more detailed data sets. If these are available they can be the most accurate indication of supply. Where they are measured from utilisation information they can be subject to bias.

Using utilisation data as a source of supply figures requires careful consideration. The number of unique health providers supplying services can often be determined from utilisation data that contains some type of provider identifier. Not all providers work full days, and such a method of simply counting unique providers makes an assumption that all providers work equivalent hours between geographic regions. Such an assumption doesn't hold well for either hospital or primary care settings. In a hospital setting providers can work in private practices or may split their time between clinic and surgery duties. In primary

care many general practitioners have variable patient contact hours depending on their other commitments. Using such data sets can result in demand bias where there is a reservoir of unused resource that is not called upon until demand reaches a certain level. Such bias can under-estimate the supply of the resource possible.

Using rosters or Full Time Equivalence (FTE) can eliminate the problems of using routinely collected utilisation data to determine service supply, but may be more difficult to source (Luo and Qi, 2009). The number of available hospital beds is a good proxy measure of available resource but only applies for in-patient services (Kalogirou and Foley, 2006; Cinnamon et al., 2008; Congdon, 2010).

There are two ways in which a PTP ratio can be expressed; as a ratio of provider:population or population:provider. The latter is usually more intuitive because population sizes are almost always greater than the number of supply units. This means that the ratio can be expressed as the number of patients per provider (e.g. 2,000 patients per provider). This allows a reader to rationalise the demand that may be placed upon any provider to supply services. The former is usually expressed as a decimal and can be less intuitive (e.g. 0.005 providers per patient). To give such a ratio more meaning it can often be expressed in terms per 100,000 patients (e.g. 500 providers per 100,000 patients). Although such an expression is easier to understand, it doesn't provide as much intuition to the demand supply scenario because neither unit is one. Expressing a PTP ratio using the provider as the numerator results in the measure correlating directly to accessibility. The lower the ratio value the lower the accessibility. Using patients as the numerator results in an inverse correlation.

Accessibility is expressed as a uniform value across the geographic area included in the calculation. This method is therefore a very gross indicator and cannot detect any change below the single geographic level that is being measured.

I used a granular measure of supply for general practice by using mean

provider days as the provider unit. The ratio was sensitive to changes in the number of days worked by GPs. I had no reliable way of determining the resources available in secondary care, and using a unit measure of facilities would have showed no difference over time. I did not therefore calculate a PTP ratio for secondary care services. It would be possible with an expanded data set that provided an ability to measure secondary care resources available.

This accessibility measure can only represent the effect the change in travel time has on utilisation of health services in a limited manner. It uses only service supply, demand and a single geographical catchment. The geographical catchment is usually defined by geopolitical boundaries and would be unlikely to change as a result of an event such as the Manawatu Gorge closure. A change in travel time can therefore only affect the measure by a subsequent change in either supply or demand. Supply could conceivably change as a response by health services to counter any perceived effect of an event that increases travel times to those services. This may take the form of opening new temporary service sites or by increasing more services in existing locations in areas affected. Demand may also change where a travel time increase is perceived as a large enough barrier to access so as to cause people to relocate their domiciles to an area not affected. This measurement is not the most appropriate for determining any changes in potential access at a sub-DHB level because there are less clearly defined geographical boundaries at a lower level than the DHB and because any such boundaries such as Territorial Local Authorities are likely to be crossed by patients seeking care.

There was little overall effect on accessibility of general practice services during the gorge closure. The end result shows that the amount of provider resource from the beginning of the study period to the end of the period increased greater than the patient population resulting in a net increase in accessibility.

The ratio however decreased immediately before the gorge was open from 7.2×10^{-04} to 7.0×10^{-04} when it opened again. This appears

to have been caused by a decrease in supply through provider resource while the patient population remained relatively constant. What caused this decrease in provide supply is not known. The timing suggests that there were likely other factors that influenced this other than just the gorge closure, but some effect may be attributable to this.

It is possible that there may have been some impact on health care workers. Any GP who lived on the other side of the gorge to where they worked may have not been available for as many days of work as previously. The mechanisms by which this may have happened are unclear. Without domicile information on providers it would be difficult to calculate the impact travel time has on the workforce.

7.5.3 Catchments

The catchment method is relatively straightforward for assessing population access to services. It involves identifying service locations, determining an appropriate catchment size or sizes and calculating the proportion of a population that resides inside each catchments. Catchment sizes can be either based on Euclidean measures or road-networks. Travel time is an often used metric in determination of catchment sizes in the health literature (Christie and Fone, 2003; Cinnamon et al., 2008). Travel distance can also be used (Lin et al., 2002; Turnbull et al., 2008).

This method requires the use of sometimes arbitrary catchment areas. A catchment area is defined by travel time or distance. Such boundaries must be chosen with care because measured accessibility drops to zero immediately outside the boundary of the catchment. Determining the appropriate travel time can be done through identifying appropriate health policy documentation, or through population sensitivity analysis as seen in Cinnamon et al. (2008).

Sensitivity analysis examines the proportion of the population included with each increasing catchment size. The appropriate catchment size

can be determined when the sensitivity value falls below a certain threshold.

Primary care services are contracted to provide first level services so that 95% of their funded patients are within 30 minutes driving time during normal business hours and are within 60 minutes out of normal business hours (District Health Board Shared Services, 2013). There are no clear reasons given why 30 and 60 minutes were chosen as the threshold for service provision. Cinnamon et al. (2008) has observed that sixty minutes may not be an appropriate travel-time threshold for access to palliative care services. They surmise that sixty minutes is often chosen due to the 'golden hour' guideline for emergency services, where patients should receive hospital care within this time to minimise chance of mortality or other serious consequences. Haynes et al. (2003) showed that patients nominate their general practitioner to be within 40 minutes drive of their domicile. This is close to the 30 minute target set in current New Zealand policy for primary care. A large proportion of 24-hour cover for general practice is now provided through emergency departments through co-located or contracted arrangements (Wilson, 2005).

Some general practices operate accident and medical services. These run hours extended beyond normal business hours but they do not provide 24 hour coverage. I have not included these services in the out-of-hours analysis because to meet the PHO contract they must provide 24 hours care.

An analysis of catchments in MidCentral was completed using a sensitivity analysis to understand at what travel time 95% of the population live from normal and after-hours health services. This was measured prior to the Manawatu Gorge closing and while it was closed. Population figures, including domicile location were determined from PHO registers. Census data for 2012 was over six years out-of-date, and PHO registers provided an ability to make point calculations for individuals rather than those at a gross aggregated meshblock level.

The results of this analysis showed that in July and October 2009 over 95% of the population were within 30 minutes drive of general practice. In January 2010 only 93% of the population were covered. An additional ten minutes of travel time were needed to cover the required population.

This substantial increase from 30 to 40 minutes is likely an artifact caused by the merger of three rural general practices to form one super-practice. The way in which PHO registers are processed has effectively meant that from January 2010 it appears as though two practices have closed. This is not strictly true. Those practices still see patients, but are considered satellite clinics. Because this analysis is done on PHO registers this decrease in access is likely an artifact rather than a real trend. Although it would be possible to create 'virtual' practices and extrapolate patient membership of these practices from past registers there would be a degree of error introduced in doing so. I chose not to undertake such a process to try and correct for this anomaly but rather understand the reason it occurred and the impact it has on the final metric.

After-hours performance as measured to the hospital is slightly lower than the goal of 95%. At 60 minutes, the coverage is approximately 93% requiring an additional 7 minutes of travel time to reach the threshold. This results in an additional 12% increase in travel time to increase coverage by 2%. This suggests a very low population density at the limits of the catchment.

The gorge closure decreased coverage by 1%. The Manawatu Gorge closing has added 5-15 minutes travel time to a trip from the eastern areas of MidCentral to Palmerston North hospital. This effectively shrinks the size of the catchment in the east by 5-15 minutes. The effect of 1% seems low, but the population density at the periphery of the catchment is very low with it being a rural area. The metric does not take into account that the closure has increased patient travel time to the hospital as long as patients are still within the catchment threshold. It does not express degrees of accessibility for populations.

People living five minutes from a service are considered to have similar access to those living at fifty-five minutes.

The catchment accessibility method is ideally suited to measure performance against the PHO contract for general practice. In this instance it shows that general practice in MidCentral is likely meeting its obligations to provide first level services to at least 95% of the population within the prescribed 30 minutes. When considering the metric prior to October 2010 the increase in time to 40 minutes is likely an artifact of the way PHO registers represent practices. The effect of the Manawatu Gorge closing on this metric was minimal. This result seems reasonable considering that a number of general practices operate on the eastern side of the gorge and anyone living there and attending a general practice there would be ultimately unaffected by the gorge closure.

It appears as though MidCentral may not be meeting their obligations for the provision of after hours services.

Catchment models do not take into account the magnitude of any supply component unlike the PTP ratio method. Each facility in an analysis is assumed to have equal supply. This results in this metric being an indicator of patients ability to access a particular service physically, but does not make any assumptions on the site's ability to deliver such services as needed by the catchment population.

The measure is a relatively intuitive one, being able to be expressed in basic terms as a percentage of the population being within a given travel time of services.

This metric is useful in assessing PHOs' conformance to existing health policy. It provides some insight into the impact on populations with regard to prescribe population coverage policies. It does not provide any insight into the impact on those that are still deemed to be within a reasonable travel time of services. It is able to illustrate a change in access to services during the Manawatu Gorge closure.

7.5.4 *Two-Step-Floating Catchments*

The Two-Step Floating Catchment Area (2SFCA) method is an improvement on other models that define single catchment areas (Luo and Qi, 2009). It accounts for the location of both providers and patients (Wang and Luo, 2005). It represents much finer scales particularly for rural populations (McGrail and Humphreys, 2009). It uses a two step process to first calculate provider-to-patient ratios for geographic areas within a constant catchment size. It then sums the ratios for each geographic area to each service site within the catchment distance.

Accessibility is expressed as an index value. Increasing index values indicate increasing access to health services.

This process eliminates some of the problems seen in the provider-to-patient ratio (McGrail, 2012). The provider-to-patient ratio calculations are often only applied to large geographical units because of the issue with zero access beyond the catchment area. In the two-step floating approach, an area's access to services is a function of the number of services available within a certain catchment distance. This results in a more granular view of accessibility of health services.

The two-step floating catchment method is not appropriate where there is only one service location. In MidCentral I am considering access to ED services for which there is only one hospital. The 2SFCA method is not therefore appropriate in my analysis of hospital services.

The main consideration in applying the 2SFCA method is what size catchment to use (McGrail, 2012). The catchment size will depend on what one is investigating. When measuring performance of PHOs against contract guidelines, catchments of 30 minutes and 60 minutes are appropriate for measuring access to normal hours and after-hours services respectively.

A catchment may be defined in terms of time or distance and may use a Euclidean or road-network measure. Small geographical units are usually used to aggregate populations and perform analysis. Mesh-

blocks are a convenient unit to use for aggregation. Although analysis results are aggregated, I made initial calculations using point coordinates. This is different to most other studies which have begun by using aggregated data to make calculations. As previously discussed, this may lead to bias when comparing rural and urban populations. It is ultimately necessary to aggregate any results into geographical areas for presentation, so travel times from individual patient calculations are aggregated by meshblock.

The areas around Dannevirke with the gorge open show moderate accessibility. This is likely in part due to a large medical centre in Dannevirke.

Some meshblocks had extraordinarily high accessibility indices up to 838. These are artifacts of extremely low population counts in some meshblocks particularly in and around parks and industrial areas in urban centres. Some meshblock mean counts were as low as 0.1. This represents meshblocks that had a handful of people residing in them for only a portion of the study period. Meshblock populations lower than 1 have a multiplicative rather than divisive effect on the index, providing such extreme values.

There are three factors that contribute to the changes in accessibility indices. The two-step floating catchment method is most sensitive to provider numbers and to a lesser extent populations. This higher sensitivity is due to provider numbers being typically much lower than population numbers, so a unit change in provider numbers is proportionately much greater than unit change in population. Over the period of the gorge closure, both of these figures fluctuated as would be expected in most real-world situations. These fluctuations however were not substantial over a large area (the aggregates of them can be seen in Table 6.5). The final factor that contributes to the change in accessibility is the increase in drive time caused by the gorge closure. This of course is the factor that this thesis is interested in addressing.

The large cluster of decreased access in the north-east is therefore likely

to be attributable to the closure of the gorge and its impact on the travel times into the eastern areas of MidCentral on those living there. This method continues to rely on catchments and as such shows differences in accessibility for those at the peripheries of those areas. It provides a more complex view of how patients may choose to access general practice services than a simple catchment based method.

The two-step floating catchment method has substantial promise as a measure for potential access to general practice services in New Zealand. It can easily account for population and provider density. It can be sensitive to very low population numbers in meshblock sized geographies. Such sensitivity may produce noise in analysis however. Using some unit larger than meshblocks may help to smooth these variances. Because the 2SFCA uses an index, it is difficult to relate its measure back to health policy documents that still refer to catchment and proportions of populations.

The two-step method modifies the basic catchment method only where some service sites have overlapping catchments. It would therefore only be useful in analysis of hospital services were there is more than one site providing services within a catchment of each other.

7.5.5 Two-Step-Floating Catchments with Distance Decay

The two-step floating catchment with distance decay (2SFCDD) method uses a process almost identical to the 2SFCA method but with the addition of a weighting applied to each population-service index (Luo and Qi, 2009). The weighting is a function of the population's distance from the service location. The further away the population is the lower the weight applied. This has the effect of decaying the accessibility index as populations move further from services. The rate at which this decay occurs is applied as an arbitrary function. Its only property needs to be that populations further from services contribute to the accessibility index less than those that are close.

Arbitrary catchment sizes must be determined for the 2SFCDD method.

The concept of catchment size is relatively intuitive and can at times be attributed to specific requirements or scenarios. The decay function is not as intuitive a criteria. How much less should a population's accessibility be indexed when they live twice as far away as another population? One method that has been presented is the 'gravity' or 'inverse square' decay function (Weibull, 1976; Joseph and Bantock, 1982; Guagliardo, 2004; Congdon, 2010).

Haynes et al. (2006) challenges the traditional gravity model approach of using an inverse square as the decay function for the gravity model. The inverse square function is a power function, and exhibits properties that produce significantly higher values for units less than one. This can overestimate the the impact of living in close proximity to health services. It also results in the function relying heavily on the units used in the decay function (i.e. metres versus kilometres).

There are ways of countering Haynes et al.'s (2006) concerns (Luo and Qi, 2009; McGrail, 2012). The problem of choosing an appropriate unit for calculations can be overcome by using proportionate distances. The distance of any population from a service can be expressed as a proportion of the total catchment size. Elimination of significantly high values can be achieved through using a minimum unit distance to apply any decay to, with all values below this minimum receiving a weighting of 100%.

Haynes et al. (2006) suggest that a decay function using an inverse exponential may be more appropriate for comparison of GP registration data compared with travel time for each registrant to their chosen surgery. They showed a very high coefficient of determination using empirical data from GP registers and road network travel time. It has been recommended that local empirical data is used to calibrate the decay function for local use (Huff, 2000).

The 2SFCDD shows lower accessibility in and around Dannevirke than the 2SFCA. The decreased effect of services at the edge of population catchments are the likely explanation for this. The population around

Dannevirke would have access to practices in Woodville and Pahiatua within the 30 minute catchment. These would likely be at the edge of the catchment area. The decay function is decreasing their influence on the accessibility index.

The change in accessibility during the gorge closure using the 2SFCDD was markedly different from the results using the 2SFCA method. The 2SFCA method showed a substantial decrease in access around Dannevirke, Woodville and Pahiatua. The 2SFCDD method shows a much less uniform pattern in Dannevirke and Woodville.

The variations are likely due to increases or decreases in the populations in those meshblocks. This effect would have been small because the gorge closure affected practices at the peripheries of these population catchments. The effect seen in and around Dannevirke is similar to the variation seen in the unaffected western areas of the district. It is therefore likely to be the effect of random variation rather than the gorge closure.

The pattern of decreased accessibility seen around Pahiatua is more uniform. Such uniformity is less likely to be the result of random variation. It indicates a prediction of decreased potential accessibility. It is the only uniform pattern that this method predicts across the district.

The explanation for the different sensitivity of the 2SFCDD and 2SFCA to the gorge closure lies with the practices at the periphery of the population catchments. Practices in Ashhurst and Palmerston North that were previously at the periphery of population catchments were no longer within those catchments when the gorge was closed. In the 2SFCDD method, peripheral practices had a minimal influence on the accessibility index when the gorge was open. In the 2SFCA method peripheral practices contributed as much to the accessibility index as practices much closer to the eastern populations. The way that the 2SFCDD behaves in predicting accessibility is likely to reflect reality where very few individuals living in Dannevirke are using general prac-

tice services in Ashhurst or Palmerston North.

The 2SFCDD appears to be a good model to reflect potential accessibility to general practice in MidCentral. It represents a slightly lower accessibility for more remote populations which could be argued reflects reality. This is consistent with international findings using a 2SFCDD to measure accessibility to primary care providers (Luo and Qi, 2009; McGrail, 2012). The number of patients who had GP access affected by the gorge closure was small (1,071). The 2SFCDD therefore presents a more balanced view of how the gorge closure affected populations access to practices at the peripheries of the catchments.

7.5.6 Summary of Potential Access

There are further variations on the 2SFCA beyond adding a decay function. These include varying the catchment sizes dependent on the type and location of services (McGrail, 2012; Yang et al., 2006), and using asymmetric catchment sizes for providers and population (Tiwari and Rushton, 2005). Although there is evidence that these can counter some of the issues inherent in the 2SFCA method, they involve making further arbitrary decision on catchment sizes. Presently in New Zealand catchment sizes of 30 and 60 minutes correspond to health policy. Applying catchments of alternative sizes would require substantial consideration of what was trying to be achieved and investigation of catchment size effects.

Travel time isochrones are a good visual aide to understand the area impact of travel time changes to a single location. They do not account for service provision. This accessibility model illustrated clearly an area that was affected by the gorge closure for service to Palmerston North hospital.

The Provider-To-Population ratios put supply and demand into context. They do not account for geography within the catchment they operate on. In this case they have suggested that a decrease in supply has impacted general practice services across the district. Few other

studies have considered how spatial barriers impact the health workforce.

The use of catchments with sensitivity analysis have a direct relationship to measuring performance against existing health policy. Catchment methodologies are intuitive. They provide a gross figure on accessibility but no insight into accessibility of a sub-catchment area. In this case it demonstrated a decrease in accessibility in both general practice and hospital services but only to a small degree.

The two-step floating catchment area method is less intuitive as it produces an arbitrary index value. It may be more useful to classify indexes into quantiles representing accessibility categories, particularly to visualise results. The 2SFCA method is highly sensitive to population and services at the peripheries of catchments. This method demonstrated a decrease in accessibility of general practice services during the Manawatu Gorge closure. This appears to be overstated when compared to the relatively modest impact the catchment sensitivity analysis suggested.

The two-step floating catchment area method with distance decay suffers from a similar problem of being unintuitive as a measure. It is less sensitive to populations and services at the peripheries of catchments. It shows a decrease in accessibility but in a more isolated geographical area than the 2SFCA method. This is more consistent with results from the catchment sensitivity analysis than the 2SFCA method and it provides additional geographical information about areas below the catchment level that the catchment method does not.

7.6 *Realised Accessibility*

The four main service areas being investigated are Emergency Department, General Practice, Hospital Admissions and Hospital Outpatients. Each of these areas exhibit different overall utilisation rates. General practice utilisation is the highest. This is to be expected because most

people attend their GP for minor to moderate ailments, general health checks and screening activities such as immunisation and cancer screening. Outpatient attendances are the next most frequently used service area. This is expected given that a single hospital admission may lead to several outpatient follow-ups. Additional to this, a number of patients may attend outpatient clinics for public specialist care with no associated hospital admission. Emergency Department attendances are much rarer than either GP or Outpatient service use, and slightly more common overall than hospital admissions. Hospital Admissions can be classified into two classes, acute and elective. Acute admissions are those that are not planned, as a result of an unexpected illness. Such admissions are often related to ED attendance, where ED are not able to resolve a problem for a patient such patients are admitted to the hospital. Elective admissions are those that are planned and not generally time critical. These may include elective surgery such as total hip replacements.

All utilisation rates for realised accessibility calculations were standardised against the funded PHO population. I was unable to source Census 2006 data in appropriate categories to standardise it at a national level. Because utilisation was clearly affected by not only age but also deprivation, ethnicity and gender I needed reference data to standardise by appropriate categories. Statistics NZ did not supply data in an appropriate way. It was possible using the low level data I had from the PHO registers to categorise into appropriate reference groupings. The outcome of standardising against local PHO registers is that the standardised utilisation rates will may not be comparable with other national data. This had no impact on the study's ability to investigate the questions at hand.

Increasing utilisation rates in older groups is in fitting with our expectation of the aging process. As people age, they will begin to have more frequent health concerns in higher proportions. They will visit the GP more often to help them manage their health, they will turn up to ED more often with acute exacerbation of conditions, and be

admitted to hospital more often for acute or elective procedures, with a consequently higher rate of outpatient attendance.

There is a minor peak in hospital admissions in the age group from 20-40 . This is a distinct trend not mirrored in any of the other services. This age range coincides with the main child-bearing years for females, and is likely the consequence of hospital births, which are recorded as hospital admissions.

The GP, Hospital Admissions and Outpatients have remarkably similar utilisation rates by ethnicity. New Zealand Europeans exhibit the highest rate of use of all three service areas. This type of effect is likely caused by the significantly different age profile that NZ Europeans have compared with NZ Maori and Pacific ethnicities (Figure 6.13). Because the proportion of those over 50 for NZ European's is much higher than the NZ Maori or Pacific and because utilisation rates are significantly higher in the over 50 age group there is a much higher overall utilisation rate when grouped by ethnicity. The effect being seen here is partly due to a comparison of an older population of people with a much younger one. It is unclear what factors are contributing to the much lower utilisation rate for the 'Other' ethnicity group, who have a similar age profile to the NZ European group.

This effect may be explicable by the different deprivation profiles seen in each ethnic group (Figure 6.9 on page 103). Both NZ Maori and Pacific ethnicities have most of their populations in the most deprived socioeconomic groups, while NZ European and 'Other' have a comparatively even distribution. Socioeconomic deprivation can impact upon use of ED services in a multitude of ways. There is some evidence that because ED services are free, some lower socioeconomic groups favour attending ED rather than their GP where patients are still expected to pay for part of their consultation costs (Hanania et al., 1997). It is also well known that living conditions contribute to the health status of individuals (Howden-Chapman et al., 2007, 2008). Those living in poorly insulated and poorly heated houses, and those that live in high resident density domiciles all exhibit higher rates of ill-health. Each of

these factors is more likely seen in the lower socioeconomic groups.

The pattern of utilisation by deprivation is likely to be demonstrating the increased health needs of those in higher deprivation groups. These increased health needs may be bought on by associated factors such as poor living conditions, poor diet, or inability manage existing health conditions. This is further supported by the profile for GP utilisation, showing the quintiles from 2-5 exhibit almost identical utilisation, with the 1st quintile showing a slightly lower rate. General practice in New Zealand still requires users of the service to meet a co-payment for using the service. This represents a financial barrier for lower socioeconomic groups. Low health literacy may result in lower socioeconomic groups from seeking care until emergency care is needed.

Sporting injuries and motor vehicle accidents may be the cause of more males attending ED. Child bearing may have an impact on more females being admitted to hospital and seeing their GP. Of the four demographic factors considered here, gender displays the least variation.

All four demographic factors confound utilisation. Therefore it is important to standardise utilisation rates, particularly for comparison between control and intervention groups.

7.6.1 Choice

In New Zealand patients have the right and ability in most situations to choose their general practice (Hays et al., 1990). A patient's right to choose their service provider is implied by health & disability law in New Zealand (Health & Disability Commissioner, 2009) and explicitly stated by some DHBs. When a practice decides that they have the most patients that they can service given the resources they have they may stop taking patients onto their register. While a practice is in this situation they may turn patients away. This represents the most common scenario where a patient may not express their choice of health service provider. There is currently no evidence to suggest

how frequently this prevents patients from their first choice of general practice.

There is an assumption in most studies that people seek care from the nearest service provider to where they live. Calculations are then made on that basis. There are a number of factors that influence the choice of where a person may seek their care. The distance to their service provider will certainly be one of those factors, however, where there is a choice of service provider within the same area, the effect of this is less clear. There are matters of the personalities of service providers. Where a patient has choice, they may choose to visit a service provider further than their closest because of interpersonal relationships. They may not move service providers if they move house, particularly in areas where it is difficult to find new service providers. They may seek a provider closer to their place of work, rather than their domicile, or seek a provider closer to their children's school. All of these factors influence a person's daily activity space. The larger a person's daily activity space, the more significant the difference may be between where they live and where they seek service.

Not having a choice in service provider may present barriers to access due to privacy issues, personality issues, or financial or debt issues. There is some suggestion that patients may not seek medical treatment for sensitive issues from their usual service provider. Having a limited choice of providers may inhibit access for sensitive issues. Only having one service provider available may present issues where there are personality issues, particularly in practices with only one doctor or nurse. In larger practices this is unlikely to be an issue. There may be cultural or ethnic issues where there are not choice of service provider. Where patients run up financial debt in a practice, they may be reluctant to attend that practice until such time as the debt is cleared, rather than when they need to. In situations where they have a choice of service provider, they may be able to seek treatment elsewhere while the debt is outstanding at their usual provider.

Using practice choice as an accessibility measure is slightly different

from the previously mentioned methods. These have all been potential accessibility measures, not accounting for actual behaviour by patients. This measure is a mix of potential accessibility and realised behaviour through nomination of a preferred practice by patients. It does not account for actual access to health care in practices.

Practice choice was calculated from patient registers and meshblock data. For each meshblock's population of resident patients, the total number of unique practices was summed. Only practices that had more than five patients in each meshblock were counted to reduce anomalies caused by natural minor variations. Often people will move house but not immediately change their general practice.

Practice choice appears to correlate well geographically with other potential accessibility measures having least accessibility in the north and eastern areas of the district. It shows a more homogenous profile through the eastern areas than the other accessibility measures, perhaps not accounting for distance as well as the 2SFCA or 2SFCDD methods which showed a difference between the far east and mid-east areas of the district.

There is a very strong relationship between maximum deprivation and practice choice ($r^2 = 0.854$). This is a relationship not seen in the other accessibility measures. By comparison the 2SFCA correlation between accessibility index and deprivation is very weak ($r^2 = 0.2249$).

Deprivation affects health status, but health status does not affect deprivation. This is due to the way deprivation is calculated. The New Zealand Deprivation Index 2006 is a composite measure based on income, home ownership, parental support, employment, qualifications, living space, telephone access, and access to a car (Salmond et al., 2007). Aspects of income, home ownership, qualifications, living space and car access may all directly affect a person's ability to seek health services. Improving a person's health state may make them more likely to seek employment, which in turn may improve their income, living arrangements and ability to afford and own a car. Any short to medium

term effect on these factors due to improved health state would be unlikely however. For this reason, it is unlikely that patient choice has a protective effect on deprivation.

The inverse care law states that the provision of health services when allowed to be driven by market forces is least in the areas where it is needed most and vice versa (Hart, 1971). General practice in New Zealand is provided on a mixed model of publicly subsidised and privately funded business. Hospital care is provided in the majority from a publicly funded system, with some elective care provided by private hospitals. The private component of both primary and secondary care mean that health care is susceptible to market forces. Having an inverse care situation is undesirable in a society where there are recognised inequalities in health outcomes such as in New Zealand (Harris et al., 2006). The aim of a number of current health policies is to modify or intervene in the market forces to reverse this effect. The inverse care law continues to pervade both developed and undeveloped countries today (Graves, 2009).

Some evidence suggests that this law may not hold for the most deprived populations (Christie and Fone, 2003). They often live in high density urban areas, in the same place major hospital services are often located. The evidence presented by the high correlation between practice choice and deprivation score suggests the contrary. In Mid-Central the association between these two factors may be an indicator that general practices will tend to operate in areas that can support them financially; not the areas that necessarily need improved access to health care.

The implication of this is if practice choice is used as a geospatial measure it is important to consider its relationship with deprivation, and any effects that deprivation may have on the factors being measured. This confounding should be controlled for.

The relationship practice choice appears to have with deprivation may have important implications for policy makers in the future when con-

sidering the support of new general practices and their location.

Practice choice is only a useful measure where patients have a choice of health facility. Where a district is served by one main facility, such as the hospital in Palmerston North, there is no patient choice to measure. In order for choice to be separate from utilisation there must also be a mechanism by which patients can indicate their choice. PHOs have an enrollment mechanism. Hospital services do not.

Practice choice is not sensitive to the closure of the Manawatu gorge.

7.6.2 ED Attendances

The overall crude attendance rates of 24.2 to 24.5 appears to be a reasonable figure compared with other literature. Hider et al. (2001) reported that attendance rates in Christchurch hospital were approximately 20 visits per 100 people in 1998. Hull et al. (1997) reported a range from 10.3 to 29.4 having a mean of 17.6 for UK hospitals in the early to mid-nineties. In an investigation of Australian emergency department attendance rates, Lowthian et al. (2012) compared the rates from 1998 to 2008 and found a contemporary rate of 20 in 100 attendances which had increased by a factor of 32% over the decade.

There is evidence that the ED attendance rate is inversely correlated with travel time from ED. There are distinct bands of attendance rates, with 0-4 minutes having the highest rate, with a second band from 5 - 24 minutes, a third band from 25 - 59 minutes, and a very sharp drop in rate in the 60+ minute travel time. These results are consistent with other studies showing that there is a correlation between ED attendance and travel time to ED (Magnusson, 1980; Stock, 1983; McKee et al., 1990; Jordan et al., 2004). The problem with this type of analysis is that although some demographics are controlled for by standardising the rates, it is difficult to control for factors as people that are prone to make numerous ED attendances may choose to live closer to ED and hospital. These findings are consistent with the body of literature that travel time affects use of health services.

The extreme drop in access beyond sixty minutes travel time may be explained by the catchment of the MidCentral Health district. The district is bounded by Hawke's Bay District Health board in the North, Wanganui District Health Board in the West, Capital and Coast District Health board in the South West and Wairarapa District Health Board in the South East. It is possible that there may be a situation where populations living at the extremes of these boarders may seek health services across these geopolitical boarders in an alternative district. This would result in their health service attendance behaviour not being captured in the data set of this study providing an artificially low rate in this category. There is however no evidence in the data that meshblocks bordering the MidCentral boundary have substantially different access rates than those in the interior of the district beyond sixty minutes travel time.

Health Policy and existing literature reference sixty minutes as a threshold figure for providing health services to the majority of the population (Brabyn and Skelly, 2001, 2002). This data supports this figure given the substantial drop in utilisation rates over this threshold. Health policy that has sixty minute travel times as thresholds for service provision appear to have a consistent basis in fact when considering utilisation rates.

There was no measured impact on ED attendance rates due to the Gorge closure. The negative result may mean that the statistical power of the study was insufficient to detect the real change. This is a relatively large total set of data however, comprising 355,096 people years of observations. The intervention group is much smaller than this at 29,246 people years.

ED attendances can range from significant events where people will have no choice but to seek medical attention at ED to minor presentations that could otherwise be seen at general practice or an after hours primary care medical facility. An increase in travel time would be unlikely to affect the most serious ED cases. The less serious cases however are sometimes considered to be made because of convenience;

due to reduced direct cost to patients; and because they operate 24 hours a day. ED presentations with a triage score of five and perhaps four could be considered presentations of convenience that may well be dealt with more appropriately in other care settings.

Presentations driven by convenience may be impacted by increased travel time where the inconvenience of the additional travel out-weights other considerations.

Presentations with a triage score of five only showed a decrease in the control and no change in the intervention group. Such a decrease suggests that there were likely other extraneous factors that contributed to non-emergency presentations to ED decreasing in this time-frame. One such factor could have been the introduction of free visits to GP after hours services for those under 6 in October 2011 (Ruscoe and Haran, 2013), right at the beginning of the intervention period of the study.

Providing free services in primary care for children after hours could conceivably affect the presentation rates to ED for ailments deemed insignificant and triaged in the lowest urgency category by removing cost of presentation as a motivator. This however wouldn't explain why there was not an associated drop in presentation rates for those affected by the gorge closure.

7.6.3 Outpatient Attendances

The crude outpatient appointment rates (Table 6.10 on page 133) are comparable between funded and study cohort and across years. I was not able to find any literature that could compare the crude rate in the data provided here with New Zealand or Australasian data. Most literature reports outpatient DNA rates but not appointment rates. This may be because of the way that outpatients are typically classified into specialties. Analysis of crude outpatient rates grouped together may provide little interest or utility outside the type of analysis I am conducting here. The crude rate for outpatient attendances is lower

than GP attendances by a factor of about 4, and higher than hospital admission rates by a factor of approximately 5-6. This seems reasonable given that people will visit general practice more often than visit a hospital outpatient clinic and they will visit a hospital outpatient clinic more times than each admission.

Outpatient attendances analysed by five minute travel time categories (Figure 6.26 on page 127) exhibit a relationship between travel time and utilisation. This relationship appears to affect both extreme travel time categories, less than five minutes and greater than 60 minutes. Travel time to outpatient services between these times shows little statistically significant difference. The large confidence interval seen in the 55-59 minute group may be due to a lower number of population living at this distance. Figure 6.14 (page 109) shows that this travel distance is one of the few that doesn't incorporate any large urban population within its catchment.

There was a statistically significant drop in outpatient attendance in the control group and a statistically significant increase in the intervention group. A drop in outpatient attendance rates in the control group may be caused by a number of factors, including a reduction in the number of clinics provided, or a reduction in the corresponding hospital admissions leading to fewer outpatient appointments. The magnitude of the increase was relatively small however, being only 1%. The corresponding increase in the attendance rate ratio for the intervention group was in the order of 10%.

The mechanisms that may cause people with an increased travel time to attend outpatient appointments more frequently are unclear. Outpatient appointments result from follow-up from hospital admissions and also from referrals from general practice for patients to see specialists. There appeared to be no increase in the number of hospital admissions during the period of the gorge closure so follow-ups seem unlikely to be the explanation for this trend. There may have been some behavioural change in general practice as a result of the gorge closure and reduced access to the hospital. General practice may have

applied a lower threshold for referring patients for specialist assessment during this period, knowing that acute emergency care may have been less available. The idea behind referring patients for specialist care earlier is to reduce the chances of acute exacerbation of conditions, to help them be medically managed better.

This result is not consistent with other literature on the topic (Athas et al., 2000; Jones et al., 1998). It suggests that use of outpatient services increases with increased travel time. It may however be an artifact of changed behaviour by general practice, who are the gatekeepers of specialist assessment services in hospital outpatient settings. No other literature on this topic has used a natural experiment in this way to investigate this phenomenon.

The result of the intervention group analysis (Figure 6.28 on page 130) compared with the travel time category analysis (Figure 6.26 on page 127) for outpatient attendances tell different stories. The travel time category analysis is the more traditional analysis undertaken by most prior studies investigating realised access to health services. These findings are consistent with the body of literature that finds a travel-time effect to outpatient services (Fortney et al., 1999). The intervention analysis allows for the control of usually uncontrollable factors including the differences in the way rural people seek health services. Intervention analysis suggests that a travel time difference of up to 15 minutes may not be a significant factor affecting access to services. It appears as though other factors have had a bearing on the access rate for both the control and intervention groups. How travel distance affects outpatient appointments may depend on the type of service being sought. No sub-analysis was carried out on outpatient attendance by medical specialty groups.

Outpatient did-not-attend rates by travel time category suggest that those living further from Palmerston North hospital fail to attend their appointments more frequently. There is some fluctuation in rates post 30 minutes with alternate travel time groups having lower DNA rates, where some confidence intervals overlap while others don't. It is diffi-

cult to postulate a scenario or reason why DNA rates would fluctuate in this manner between five minute travel time categories. Regardless of this fluctuation the overall trend certainly appears to be a higher DNA rate in the more distant population.

Outpatient did-not-attend rate ratios dropped substantially during the gorge closure for the control group by 15%. There was no statistically significant drop for the intervention group. The p-value of 0.057 was just outside the 0.05 level set for this study. When put into perspective that during the same period there was a corresponding increase in the outpatient attendance rate of 8% (while the control group only had an increase of 1%) there is little evidence that there is any difference between the control and intervention groups. While some factor caused a drop in DNA rate across the population, it appears to have been applied to both the control and intervention groups.

There are a number of mechanisms that can be applied at a service level that may decrease the DNA rates for outpatient appointments. Increased appointment waiting times may make getting appointments more difficult. This may in turn make hospital outpatient appointments seem to be more of a scarce resource with the consequence that patients make more of an effort to attend. Increased waiting times for outpatient appointments may also have the counter-effect, where appointments are made much earlier than the appointment date, there is a higher chance patients may forget about the appointment. There is also a higher chance that they may die or move out of the area. These two latter effects are controlled for in this study as patients in the analysis who died or moved were excluded from the cohort. Hospital outpatient departments can also affect DNA rates by sending reminder letters, or text messaging to alert patients about their appointments a few days prior to them which effectively drops the DNA rate (Downer et al., 2005; Frankel et al., 1989).

Siedner et al. (2013) showed a correlation between days late to HIV clinics and travel distance by rural patients in Africa. The results seen in this thesis however do not support this finding. it is possible other

influencing factors decreased the DNA rate masking any travel time effect.

7.6.4 Hospital Admissions

The crude rate of hospital admissions for all funded patients (Table 6.11 on page 134) is similar to a previously published age standardised annual rate of 0.22 (Ministry of Health, 2012). The cohort group exhibit a lower rate than the funded population. The most substantial difference in demographics between the cohort and funded population is age (Figure 6.8 on page 101). The funded population has a higher proportion of people in the over 65 age groups. People at this age use hospital in-patient services more (Ministry of Health, 2012). This is likely to account for the difference in crude rates. This is unlikely to affect the analysis in this study as rates in both the time category and control and intervention analysis are age standardised.

The reason for the cohort exhibiting a progressively increasing admission rate from 2009 to 2012 is less clear. The rate increases by 0.01 (1%) each year. It is possible that aging of the cohort contributes to this trend because the cohort represents the same individuals year on year. Consequently the mean age of the cohort increases each year and a higher proportion of patients will be in the over 65 year old category. How much this affects the admission rate is not certain, and whether other factors could be creating this pattern is difficult to ascertain.

The pattern in hospital admission rates by travel time category is perplexing (Figure 6.26 on page 127). Up to the 30 minute travel time category there is evidence that those living further from the hospital have lower admission rates than the closest group of patients. This is then reversed up to 59 minutes and over 60 minutes there is an approximately five-fold reduction in rate. The rise in admission rate at 30-59 minutes is unexpected and the very substantial drop over sixty minutes is extreme.

The reason for the increase in admission rate for 30-59 minutes may be

explained by either the geography of the region or through how admission rates relate to GP and ED rates. The travel time isochrones map (Figure 6.14 on page 109) reveals that there are a number of urban settlements between 30-59 minutes travel time from Palmerston North hospital. One of these settlements is Levin. Some hospital in-patient services are offered in Levin but this was not accounted for in the travel time calculation to hospital services. It was not possible to determine the location of the inpatient event from the secondary care data set supplied so it is not clear if this included only Palmerston North hospital admissions or those including the Horowhenua Health Centre. Levin lies within the 50-55 minute travel time catchment from Palmerston North hospital. This could explain the high admission rates in this category and why it decreases moving back to approximately 30 minutes. This time is just prior to the halfway point between Palmerston North and Levin.

An alternative reason for the high admission rate in the 30-59 minute travel time category could be a relationship with reduced access rates in preventative care provided in general practice and followup outpatient services. There is a statistically significant drop in GP access rates beyond 55 minutes but no such associated drop in rates for outpatient attendance in this data. If access to these services prevent people from being hospitalised it stands to reason that there may be an inverse relationship between these measures.

The admission rate at travel times beyond sixty minutes is consistent with similar drops seen in the other utilisation measures. Hospital admission rate had the most extreme drop of all utilisation rates in the over sixty minute drive time category. The exact reason that this may occur is unclear. It is possible that there is a self-selection effect taking place. The majority of the area that is greater than sixty minutes travel from Palmerston North hospital is extremely rural in the northern and western areas of district. Hospital admissions are often an indication of a significant health event, usually as the result of a serious health issue. It is possible that people with significant health issues deliberately

choose not to live in an area that is remote.

There was no significant change in the hospital admission rate ratio for either the control or intervention group. This pattern is consistent with that seen in the ED data. It suggests that the changes seen in the utilisation rate by travel time category may be being influenced by factors other than travel time. It may alternatively suggest that up to a fifteen minute change in travel time was not a substantial change in travel time for those living beyond twenty minutes drive to health services normally.

Hospital admissions results from elective or acute situations. Acute presentations are driven by ED attendance. As there was no change in ED attendance rate ratios this result is consistent. Elective admissions are planned and not urgent. There may be some protection factors at work with elective hospital admissions. Because admissions will generally be for more significant health concerns it is possible patients' willingness to travel further is increased. Because admissions last a matter of days it is possible that a small increase in travel time being a matter of minutes is insignificant as a proportion of total appointment time. Increased travel times may seem worse when the increase in travel time approaches the length of the appointment as it may in general practice presentations.

Case complexity of hospital admissions was measured using a proxy of case-weighted discharge scores. Those in the intervention group had a higher case complexity than the control group prior to the gorge closure (Table 6.12 on page 135). The control group also experienced an increase in case-complexity during the closure.

The higher case-complexity of the intervention group compared with the control group is consistent with the idea that those that have less access to health services will ultimately present with more severe illness. Travel time category analysis confirms that those living further from both the hospital and general practice use health services less. The intervention group live further from the hospital and general prac-

tice (Table 6.4 on page 100) . The intervention group therefore uses health services less than the control group. The statistically significant higher case-weighted discharge score suggests that the intervention group present to hospital with more severe or complex disease.

The increase in case complexity during the gorge closure in the control group was not seen to the same degree in the intervention group. This suggests that there was some factor influencing case complexity in the eastern areas of the district, or that was influencing the whole district but had less of an effect in the intervention group. It is unclear what this may have been.

7.6.5 GP Attendances

The crude rate of GP Attendance was consistent with the literature ranging from 4.3 consults per patient per year to 4.6 across both the patient population and cohort groups (Table 6.13 on page 136). Rates in the UK have been reported to range between 3.5 and 5.3 (Fleming, 1989; Hippisley-Cox et al., 2007). The cohort group exhibits a slightly higher consultation rate than the population. It is possible that this is related to the younger group of people excluded from the cohort (Figure 6.8 on page 101) who would seek care less often (Figure 6.20 on page 120). This is unlikely to have had any significant impact on the final accessibility analysis because both intervention and control groups are drawn from the cohort group of patients, and all rates were age standardised for comparison.

GP utilisation rate by travel time category (Figure 6.26 on page 127) has a downward trend with increasing travel time. The pattern of GP utilisation is very close that to that seen for secondary care outpatient attendances.

Both the control and intervention groups experienced a decrease in realised accessibility during the gorge closure with p-values less than 0.001 in both cases (Table 6.8 on page 129). This suggests that there were factors at work that impacted the wider health system in the

district. The provider-to-population ratio suggests that there was a lower supply of health providers across the district during the gorge closure. This may have been a contributing factor to this trend. With fewer providers it may have been more difficult for people to attain timely general practice appointments. If general practice in areas were already running at capacity prior to the supply drop a decrease in supply would impact overall utilisation rates.

The decrease in accessibility the intervention group experienced was greater than the control group during the gorge closure. The rate ratio confidence intervals for the control and intervention group do not overlap for GP utilisation (control upper CI = 1.031, intervention lower CI = 1.055). This indicates a statistically significant difference in the magnitude of the change. This may be an indicator that travel time has had an additional burden on the intervention group, over and above the effect of the burden seen in the control group. It is also possible that whatever factor affected the control group had more of an effect on the intervention group. It is difficult to imagine what this may be. In the absence of any other evidence the most likely explanation is therefore that travel time has caused an additional burden on the intervention group.

Patients have a choice of which general practice that they enroll with, unlike attending hospital services, where in MidCentral there is only one hospital. Patients also have a choice of general practices in the eastern area of MidCentral. Because of this, the vast majority of those that live in the eastern region of the district are also enrolled in the eastern practices. This results in a very low number of patients that experienced a travel time increase from their domicile to their general practice. Only 1,071 patients were enrolled in a general practice requiring them to use the gorge during the study period compared with 8,888 that were affected in their use of hospital services. Caution must therefore be made in drawing conclusions from this result because of the small size of the intervention group.

General practice is the easiest of the utilisation measures for patients to

influence. General Practice will mostly see patients with acute minor ailments, or more serious conditions for management and ongoing monitoring. Delaying seeking treatment by patients for such conditions will be unlikely to have any immediate health impact on them. Patients can self book appointments which also remove barriers to them choosing when to present. Because general practice utilisation can be easily influenced by patient behaviour it appears sensible that this utilisation measure is the one most influenced by travel time in the intervention analysis.

7.7 Study Limitations

I implemented the Brabyn-Skelly travel-time model in a slightly different manner to that described by its original authors. It was originally described as using a derived variable to determine urban roads by using a ground cover layer. I was not able to re-create this derived variable and used Statistics New Zealand urban area information instead.

Validation of the Brabyn-Skelly model against empirical data confirmed a trend that was consistent with previous validation of this method. Given this it is likely that the model performed similarly to its original implementation. The margins of error may have been affected, but in which direction is uncertain.

Even if this minor change in implementation had some difference on the performance of the model, it has been unlikely to affect the overall accessibility results of the study. The MacRae model which I used to calculate travel times conformed well to the empirical data.

The empirical data I collected was done in a narrow band of conditions only. Because of the methodology I employed to collect traffic survey data I only obtained travel time measurements for travel in daylight hours, on sealed roads and for trips that were below twenty minutes. Collection of data was not done in dense urban environments. These factors limit the ability to draw conclusions on the models suitability

to predict travel time in these scenarios.

Horowhenua Health Centre was omitted as a service location for the inpatient and outpatient services it provides. This service was not factored into potential accessibility measures or realised travel-time category calculations. Because of its location in Levin, approximately a thirty minute drive from Palmerston North hospital the impact of omitting it as a service centre would have been to have an apparent reduction in accessibility for inpatient and outpatient services.

Omission of the Horowhenua Health centre likely affected calculations for the control group for outpatient attendance, DNAs and in-patient services. Calculations for the gorge open and closed periods would have been affected equally. It would not have impacted the intervention group calculations. It appears as though it may have had some impact on the travel-time category analysis lessening the correlation between travel time and accessibility. The centre does not provide emergency services so would not have affected calculations for these. General practice services provided out of the health centre were included in the general practice analysis. The overall impact of this omission has likely had little bearing on the overall findings of this thesis given its main impact was on the traditional travel time category analysis.

While one of the strengths of this study was its use of high fidelity PHO register data to identify patients it left the analysis vulnerable to an amalgamation of practices in the east area of the district during the study period. Three general practice sites were combined into one general practice. Patients continued to be seen at all three sites, but the PHO registers were not able to be used accurately to identify patients 'home' site. This ultimately had an impact on measures of potential accessibility for the eastern MidCentral district and was most apparent in the catchment sensitivity analysis (Table 6.6). Analysis of the change in potential accessibility during the gorge closure using 2SFCA and 2SFCDD would likely have been unaffected, as the merger of the practices and subsequent change in PHO registers occurred well before the gorge closure and this analysis was concerned with change

not overall accessibility.

In order to provide a homogenous intervention and control group the study method excluded patients who did not reside in the MidCentral district for the whole study period and where not funded in a Mid-Central PHO for the whole study period. This meant that a large proportion of children below the age of five were excluded from the study. Any effect this may have had was mitigated by ensuring calculations were age standardised. The effect that this has had on other groups of people excluded is less clear. Regardless of these exclusions the study cohort was a significant size representing 101,456 patients and 355,095 patient years.

This study made travel-time calculations on the assumption that all people have access to a motor vehicle and use this to access their health care. This is not the case. Although New Zealand has a high rate of motor vehicle ownership some people will travel to health services using public transport, by foot or bicycle. Martin et al. (2008) has observed that most studies have not incorporated public transport into accessibility modeling but that this should be a consideration as models and modeling tools improve. For the purposes of this study, the exclusion of public transport into the travel-time model has likely had some impact on the travel-time category analysis by underestimating travel time to services for those that use alternatives to private motor vehicles. It is unlikely to have had a significant impact on the intervention analysis as the assumption has been applied equally to both the control and intervention group.

Chapter VIII

Conclusion

This study used a natural experiment to investigate the impact that an increase in travel time had on accessibility to health services of a homogenous population compared with a control group. The overall findings suggest that general practice service use was decreased by small changes in travel time and that emergency department, inpatient and outpatient hospital services were not. There appeared to be other unknown factors that affected access to both general practice and hospital services during the study period. For secondary care services these effects had more of a net impact than that of any travel time effect.

The findings for hospital services were contrary to the majority of current literature on the topic. Using a more traditional approach of correlating utilisation with travel time category without a control group yielded similar results to those seen in other literature. There was a definite travel-time effect. This suggests that the correlation was unlikely to be causal. It has been previously suggested that the correlation seen in such analysis is caused by attitudinal differences in those that live rurally, or that people needing frequent health care self-select to live closer to services. The findings of this thesis would support these assertions.

By using the natural experiment of the Manawatu Gorge closure this thesis has been able to control for factors not controlled in prior literature. These factors include attitudinal differences in those that live rurally and for those that self-select their location based on need as well as changes within the health system that may affect access rates.

The gorge closure affected only part of the district. It was possible to study the utilisation patterns of those who did not experience a change in travel time at the same time as those that did. This allowed me to control for factors within the health system that may have affected access to services. By examining the change in utilisation of those affected by the gorge closure before and during the closure, I was able to control for attitudinal differences and self-selection bias.

Most prior studies have looked at different groups of people and tried to control for biases by using demographic characteristics. The results of this study suggested that the factors that influence health seeking behaviour are so complex that standard demographic profiles cannot control for all of the variation in health service utilisation.

All measures of potential access to services showed differences in accessibility to both hospital and general practice services across the district. Potential access to general practice services was much lower in rural eastern areas of MidCentral. All methods illustrated a consistent gross pattern. There were variations associated with apparent population density and proximity to urban settlements. Practice choice appeared to be the least sensitive measure to population density, followed by the two-step floating catchment area (2SFCA) and then the two-step floating catchment area with distance decay (2SFCDD) methods.

The way in which potential access measures detected accessibility change during the gorge closure is more varied. The 2SFCA showed a very consistent pattern of decreased access close to populated centres in the east. This was most likely a reflection of decreased access to general practices in Palmerston North and Ashhurst. It likely overstated the impact of accessibility on general practice services given that there were only 1,071 patients who had a travel time increase to their general practice during the gorge closure. The 2SFCDD method showed a more concentrated pattern of decreased access around Pahiatua with much larger variation across other areas. This correlated well with the minimal impact the catchment sensitivity analysis predicted during the gorge closure.

The provider-to-patient ratio analysis suggested that there was a decrease in supply of general practice services during the gorge closure. This may have been a result of increased travel time for general practice staff resulting in them taking more leave, or working fewer shifts due to the increased travel time burden. This finding has implications for service planning in longer term scenarios such as the gorge closure. Further study into the impact of travel time to place of work on health professionals may yield further interesting results.

Using a combination of both 2SFCDD and catchment sensitivity analysis can provide what appear to be sensible measures of potential accessibility measures for general practice in a New Zealand context. The catchment sensitivity analysis is directly relevant to the measurement of existing policy adherence. The 2SFCDD provides insight into accessibility at a sub-regional level.

Although general practice access decreased there appeared to be no impact on subsequent hospital admission rates. The study period ended only a few months after the gorge was reopened. This may not have been sufficient time for decreased access to General Practice to have an effect on hospital admission rates. Further follow up research of hospital admission rates past December 2012 would be warranted to investigate if decreased general practice access has any longer term effect on hospital admissions.

This study has found a previously unreported relationship between practice choice and maximum deprivation. This suggests that market forces may still influence where general practices operate in fitting with the inverse care law despite deliberate policy to counter this effect.

This study had some advantage in using PHO registers to identify patients and their place of residence. It did not rely on generalise assumptions based on census population data. Small changes in populations could be detected because of the relatively high fidelity of patient registers.

8.1 *Implications for MidCentral*

The closure of the Manawatu Gorge for 12 months was a significant event. It cost the region \$62,000 per day it was closed (Forbes, 2011). An event of this magnitude and length is rare. Specific lessons to be learned from this study are therefore limited in their application. There are wider policy implications for those living at substantial travel times from health services.

Prior literature suggests that there may be an impact on patients seeking health services who live at greater distances from them. Increased travel time may decrease their use. If this situation was indeed true it would represent an equity of access issue having the gorge closed. There was no evidence that the gorge closure impacted hospital service use by those living in the eastern areas of MidCentral. On the contrary service use increased in comparison to the control group which suggests the MidCentral DHB may have had mitigation strategies in place, or that providers and patients altered their behaviour during this time. It is likely that a combination of all three of these factors contributed to the increase in service use in some areas. While increased travel time may have caused some impedance to access to services, these other factors appeared to have more than compensated.

Whatever response MidCentral had to the gorge closure they can take comfort in the fact that they appear to have either deliberately or inadvertently mitigated the impact it may have had on hospital services. Planning for a similar response in future seems like a reasonable approach to continue with.

General Practice access was decreased in the both control and intervention groups during the gorge closure. The intervention group suffered a larger decrease in access than the control group. This suggests that there were factors at work that contributed to a system wide decrease in access to general practice during this time. There also appeared to be factors which decreased access in the intervention group at a higher rate. Across the district supply of general practice services may have

been lower during the gorge closure, which may have contributed to this decline in overall access. MidCentral may wish to consider ways in which to mitigate the decrease in access to general practice services for the small group of people affected if the gorge is closed for extended periods in the future.

This analysis confirms that there is a distinct pattern of significantly lower service utilisation of health services in almost all services for those living at great distance from them. This pattern may be caused by differences in attitudes of those that live rurally and may be the result of some self-selection in the location in which people live. These explanations are far from certain however. Further work to answer these questions may shed light on the health implications of attitudinal differences to health seeking behaviour and the long term impact it has on the health of individuals and the population as a whole. This remains relevant for MidCentral as they have a large rural catchment in the north and east of the district.

8.2 *Implications for New Zealand*

Rural aspects of New Zealand have a sparse transport network. There are numerous examples of communities that can be isolated by limited road network closures. These closures may be caused by slips, floods, earthquakes or coastal erosion.

Knowing that changes in travel times can affect access to general practice services may have implications for policy makers and service leaders. In the aftermath of any event that alters people's spatial access to health services it is important for health services to respond with initiatives that can offset the increased spatial barriers. There is evidence that this may have happened in MidCentral with some success. Other areas around the country may be able to learn lessons from MidCentral's response to the gorge closure.

Existing health policy sets out catchments of 30 and 60 minutes for

measuring health service accessibility levels for normal hours and after-hours services respectively. An analysis of service use and travel time categories supports the inclusion of a 60 minute travel time threshold. All services exhibited a statistically significant difference in service use beyond a 60 minute drive time. Even though there is evidence that the correlation between drive time and utilisation rate is not causal there appears to be some factor that is associated with those that live beyond 60 minutes from services that result in them using health services less. If policy makers want to judge the proportion of the population that are using health services at an acceptable rate then this appears to be an appropriate threshold.

The use of a 30 minute catchment to measure accessibility for day-time general practice services is less well supported by this study. There appears to be a similar appreciably drop in access rates beyond 55 minutes travel time. This would suggest that a more appropriate measure of health service coverage may be to define 55 minute catchments. To keep this measure consistent with out of hours services, it may make sense to measure catchments at sixty minutes.

This thesis has been completed using Free and Open Source Software (FOSS) on large operational data sets. This demonstrates the utility of FOSS in geospatial analysis for practical applications. I hope that this may encourage non-governmental agencies and not-for-profit organisations in New Zealand to consider this type of software stack to develop geospatial capabilities in the future.

8.3 *Contributions to the Field*

In the course of this thesis I have presented two methods to improve the data quality of geocoded domicile information from PHO registers. Any person using such a data set should check it for similar problems to those that I've found. Approximately 60% of uncoded addresses may be able to be coded using the methods I have developed.

I have validated an implementation of the Brabyn-Skelly travel time model against empirical data. This validation has confirmed that this model performs well and provides an independent assessment of the technique.

I have presented methodologies for using derived variables in travel time models that have not before been presented in the literature. I have described the use of intersection density and maximum and total turning angles and used these with success in developing a new travel time model. The new travel time model I have developed is based on NZ Open GPS data and has performed better than the incumbent model in my validation tests. Although unsuccessful with deriving road elevation as an attribute I have presented its use as a possibility for future models.

In the course of developing a new travel time model using derived variables I have also described a novel technique to allow existing least cost path analysis algorithms to account for such derived variables. I have named this technique two-step least cost path analysis. This approach allows the use of dynamic variables in algorithms that require static cost values.

Bibliography

- Andrulis, D. P. (1998). Access to care is the centerpiece in the elimination of socioeconomic disparities in health. *Annals of Internal Medicine*, 129(5):412–6.
- Astell-Burt, T., Flowerdew, R., Boyle, P. J., and Dillon, J. F. (2011). Does geographic access to primary healthcare influence the detection of hepatitis C? *Social Science & Medicine*, 72(9):1472–81.
- Athas, W. F., Adams-Cameron, M., Hunt, W. C., Amir-Fazli, A., and Key, C. R. (2000). Travel distance to radiation therapy and receipt of radiotherapy following breast-conserving surgery. *Journal of the National Cancer Institute*, 92(3):269–71.
- Auckland Regional Council (2003). Auckland regional land transport strategy 2003: Trends, impacts and causes. Technical report, Auckland Regional Council, Auckland.
- Bagheri, N., Benwell, G., and Holt, A. (2008). Modelling Accessibility to Primary Health Care using a Spatial Accessibility Index and a Need Index. *Hawai'i Journal of Public Health*, 1(1):14–27.
- Bagheri, N., Benwell, G. L., and Holt, A. (2005). Measuring spatial accessibility to primary health care Study area New Zealand health policy. In *The 17th Annual Colloquium of the Spatial Information Research Centre*, pages 2–7.
- Baker, D. W., Parker, R. M., Williams, M. V., Clark, W. S., and Nurss, J. (1997). The relationship of patient reading ability to self-reported health and use of health services. *American Journal of Public Health*, 87(6):1027–30.

- Barnett, J. R. (2000). Rationalising Hospital Services: Reflections on Hospital Restructuring and its Impacts in New Zealand. *New Zealand Geographer*, 56(1):5–21.
- Barnett, R. and Malcolm, L. (2010). Practice and ethnic variations in avoidable hospital admission rates in Christchurch, New Zealand. *Health & Place*, 16(2):199–208.
- Bateman, I., Brainard, J., and Lovett, A. (1995). Modelling Woodland Recreation Demand Using Geographical Information Systems: A Benefit Transfer Study. In *CSEERGE Working Paper GEC 95-06*, Norwich. Centre for Social and Economic Research on the Global Environment, University of East Anglia.
- Beck, R. G. (1974). The Effects of Co-Payment on the Poor. *The Journal of Human Resources*, 9(1):129–142.
- Beere, P. and Brabyn, L. (2006). Providing the evidence: Geographic accessibility of maternity units in New Zealand. *New Zealand Geographer*, 62(2):135–143.
- Beney, J., Bero, L., and Cm, B. (2009). Expanding the roles of outpatient pharmacists : effects on health services utilisation , costs , and patient outcomes. Technical Report 1, Cochrane Collaboration.
- Berki, S. E., Ashcraft, M. L. F., and Newbrander, W. C. (1984). Length-of-Stay Variations Within ICDA-8 Diagnosis-Related Groups. *Medical Care*, 22(2).
- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., and Crotty, K. (2011). Low Health Literacy and Health Outcomes : An Updated Systematic Review. *Annals of Internal Medicine*, 155(2):97–107.
- Bindman, A., Grumbach, K., Osmond, D., Komaromy, M., Vranizan, K., N, L., Billings, J., and Steward, A. (1995). Preventable Hospi-

- tilizations and Access to Health Care. *The Journal of the American Medical Association*, 274(4):305–311.
- Brabyn, L. (2002). Modelling Population Access to New Zealand 's General Practitioners. In *Proceedings of the 22nd Annual ESRI International User Conference*.
- Brabyn, L. and Barnett, R. (2004). Population need and geographical access to general practitioners in rural New Zealand. *New Zealand Medical Journal*, 117(1199):1–13.
- Brabyn, L. and Skelly, C. (2001). Geographical Access to Services , Health (GASH): Modelling Population Access to New Zealand Public Hospitals. In *The 13th Annual Colloquium of the Spatial Information Research Centre*.
- Brabyn, L. and Skelly, C. (2002). Modeling population access to New Zealand public hospitals. *International Journal of Health Geographics*, 1(1):3.
- Buerhaus, P., Auerbach, D., and Staiger, D. (2000). Implications of an aging registered nurse workforce. *Journal of the American Medical Association*, 283(22):2948–2954.
- Campbell, J. (1995). General Practitioner Appointment Systems, Patient Satisfaction, and use of Accident and Emergency Servicesa Study in One Geographical Area. *Family Practice*, 11(4):438–445.
- Christie, S. and Fone, D. (2003). Equity of access to tertiary hospitals in Wales: a travel time analysis. *Journal of Public Health*, 25(4):344–350.
- Cinnamon, J., Schuurman, N., and Crooks, V. a. (2008). A method to determine spatial access to specialized palliative care services using GIS. *BMC Health Services Research*, 8(140).

- Clark, S. and Watling, D. (2005). Modelling network travel time reliability under stochastic demand. *Transportation Research Part B: Methodological*, 39(2):119–140.
- Comber, A. J., Brunsdon, C., and Radburn, R. (2011). A spatial analysis of variations in health access: linking geography, socio-economic status and access perceptions. *International Journal of Health Geographics*, 10(1):44.
- Compass Health (2012). Compass Health Atlas 2012. Technical report, Compass Health, Wellington.
- Congdon, P. (2010). A Bayesian Approach to Prediction Using the Gravity Model, with an Application to Patient Flow Modeling. *Geographical Analysis*, 32(3):205–224.
- Considine, J., LeVasseur, S. a., and Villanueva, E. (2004). The Australasian Triage Scale: Examining emergency department nurses’ performance using computer and paper scenarios. *Annals of Emergency Medicine*, 44(5):516–523.
- Critchlow Limited (2012). Geostan Geocoding. <http://www.critchlow.co.nz/geostan/geocoding.htm> \delimeter"026E30F\$#D13. [Online. Last accessed: 3 July 2012].
- Cromley, E. K. and Shannon, G. W. (1986). Locating ambulatory medical care facilities for the elderly. *Health Services Research*, 21(4):499–514.
- Cumming, J. (2011). Integrated care in New Zealand. *International Journal of Integrated Care*, 11(Special 10th Anniversary Edition).
- Dijkstra, E. (1959). A note on two problems in connexion with graphs. *Numerishce Mathematik*, 1:269–271.

- District Health Board Shared Services (2013). PHO Services Agreement. Technical report, District Health Board Shared Services, Wellington, New Zealand.
- Downer, S. R., Meara, J. G., and Da Costa, A. C. (2005). Use of SMS text messaging to improve outpatient attendance. *The Medical Journal of Australia*, 183(7):366–8.
- Ekonomou, E., Fan, L., Buchanan, W., and Thüemmler, C. (2011). An Integrated Cloud-based Healthcare Infrastructure. In *Cloud Computing Technology and Science (CloudCom), 2011 IEEE Third International Conference on*, pages 532–536. IEEE.
- Ellison-Loschmann, L. and Pearce, N. (2006). Improving access to health care among New Zealand’s Maori population. *American Journal of Public Health*, 96(4):612–7.
- Fleming, D. M. (1989). Consultation rates in English general practice. *The Journal of the Royal College of General Practitioners*, 39(319):68–72.
- Forbes, M. (2011). Manawatu Gorge closure costs about \$62,000 a day. <http://www.stuff.co.nz/dominion-post/news/5664229/Manawatu-Gorge-closure-costs-about-62-000-a-day>. [Online. Last accessed: 23 October 2012].
- Fortney, J., Rost, K., Zhang, M., and Warren, J. (1999). The impact of geographic accessibility on the intensity and quality of depression treatment. *Medical Care*, 37(9):884–93.
- Frankel, S., Farrow, A., and West, R. (1989). Non-attendance or non-invitation? A case-control study of failed outpatient appointments. *British Medical Journal*, 298(6684):1343–5.
- Fuggetta, A. (2003). Open source softwarean evaluation. *Journal of Systems and Software*, 66(1):77–90.

- Fuller, J., Edwards, J., Procter, N., and Moss, J. (2000). How definition of mental health problems can influence help seeking in rural and remote communities. *The Australian Journal of Rural Health*, 8(3):148–53.
- Gallagher, M., Pearson, P., Drinkwater, C., and Guy, J. (2001). Managing patient demand: a qualitative study of appointment making in general practice. *The British Journal of General Practice*, 51(465):280–5.
- Gately, C., Rogers, A., and Sanders, C. (2007). Re-thinking the relationship between long-term condition self-management education and the utilisation of health services. *Social Science & Medicine*, 65(5):934–45.
- Gauld, R. (2012). New Zealand’s post-2008 health system reforms: toward re-centralization of organizational arrangements. *Health Policy*, 106(2):110–3.
- Goetzel, R. Z., Anderson, D. R., Whitmer, R. W., Ozminkowski, R. J., Dunn, R. L., Wasserman, J., and Committee, T. H. E. R. O. H. R. (1998). The Relationship Between Modifiable Health Risks and Health Care Expenditures: An Analysis of the Multi-Employer HERO Health Risk and Cost Database. *Journal of Occupational and Environmental Medicine*, 40(10).
- Goodman, D. C., Fisher, E., Stukel, T. a., and Chang, C. (1997). The distance to community medical care and the likelihood of hospitalization: is closer always better? *American Journal of Public Health*, 87(7):1144–50.
- Google (2013). The Google Geocoding API. <https://developers.google.com/maps/documentation/geocoding/#Limits>. [Online. Last accessed: 10 June 2013].

- Graves, A. (2009). A model for assessment of potential geographical accessibility : a case for GIS. *Online Journal of Rural Nursing and Health Care*, 9(1):46–55.
- Guagliardo, M. F. (2004). Spatial accessibility of primary care: concepts, methods and challenges. *International Journal of Health Geographics*, 3(1):3.
- Gulliford, M. C. (2002). Availability of primary care doctors and population health in England: is there an association? *Journal of Public Health Medicine*, 24(4):252–4.
- Hale, M. and Sharpe, N. (2011). Persistent rheumatic fever in New Zealand—a shameful indicator of child health. *The New Zealand Medical Journal*, 124(1329):6–8.
- Ham, C. (1997). Priority setting in health care: learning from international experience. *Health Policy*, 42(1):49–66.
- Hanania, N. A., David-wang, A., Kesten, S., and Chapman, K. R. (1997). Factors Associated With Emergency Department Dependence of Patients With Asthma. *Chest*, 111(2):290–5.
- Harris, R., Tobias, M., Jeff, M., Waldegrave, K., Karlsen, S., and Nazroo, J. (2006). Effects of self-reported racial discrimination and deprivation on Mori health and inequalities in New Zealand : cross-sectional study. *Lancet*, 367:2005–2009.
- Hart, J. T. (1971). The Inverse Care Law. *Lancet*, 297(7696):405–412.
- Haynes, R., Jones, A. P., Sauerzapf, V., and Zhao, H. (2006). Validation of travel times to hospital estimated by GIS. *International Journal of Health Geographics*, 5:40.
- Haynes, R., Lovett, A., and Sünnerberg, G. (2003). Potential accessibility, travel time, and consumer choice: geographical variations in

- general medical practice registrations in Eastern England. *Environment and Planning A*, 35(10):1733–1750.
- Hays, S. M., Kearns, R. a., and Moran, W. (1990). Spatial patterns of attendance at general practitioner services. *Social Science & Medicine*, 31(7):773–81.
- Health & Disability Commissioner (2009). Code of Health and Disability Services Consumers’ Rights. Technical Report 1, Office of the Health & Disability Commissioner.
- Heatwole, C. and Lohani, M. (2004). Estimating Stream Length and Sinuosity Using GIS. In *Assessing and Managing Virginia’s Water Resources*, pages 131–139.
- Hector, D., Espinel, P., and King, L. (2012). Reducing the risk of chronic disease in older adults: A summary report to support obesity prevention planning in NSW. Technical Report October 2012, Physical Activity Nutrition & Obesity Research Group, Sydney.
- Hefford, M., Crampton, P., and Foley, J. (2005). Reducing health disparities through primary care reform: the New Zealand experiment. *Health Policy*, 72(1):9–23.
- Hewitt, M. (1989). *Defining "rural" areas: Impact on health care policy and research*. DIANE Publishing.
- Hider, P., Helliwell, P., Ardagh, M., and Kirk, R. (2001). The epidemiology of emergency department attendances in Christchurch. *New Zealand Medical Journal*, 114(1129).
- Higgs, G. (2004). A Literature Review of the Use of GIS-Based Measures of Access to Health Care Services. *Health Services & Outcomes Research Methodology*, 5:119–139.

- Hippisley-Cox, J., Fenty, J., and Heaps, M. (2007). Trends in Consultation Rates in General Practice 1995 to 2006 : Analysis of the QRESEARCH database. Technical Report February, The Information Centre for Health and Social Care.
- Hiscock, R., Pearce, J., Blakely, T., and Witten, K. (2008). Is neighborhood access to health care provision associated with individual-level utilization and satisfaction? *Health Services Research*, 43(6):2183–200.
- Hodge, G. (2008). *The geography of aging: Preparing communities for the surge in seniors*. McGill-Queen’s Press-MQUP.
- Hopman, W. M., Harrison, M. B., Coe, H., Friedberg, E., Buchanan, M., and VanDenKerkhof, E. G. (2009). Associations between chronic disease, age and physical and mental health status. *Chronic Diseases in Canada*, 29(3):108–16.
- Howden-Chapman, P., Matheson, A., Crane, J., Viggers, H., Cunningham, M., Blakely, T., Cunningham, C., Woodward, A., Saville-Smith, K., O’Dea, D., Kennedy, M., Baker, M., Waipara, N., Chapman, R., and Davie, G. (2007). Effect of insulating existing houses on health inequality: cluster randomised study in the community. *British Medical Journal*, 334(7591):460.
- Howden-Chapman, P., Pierse, N., Nicholls, S., Gillespie-bennett, J., Viggers, H., and Cunningham, M. (2008). Effects of improved home heating on asthma in community dwelling children : randomised controlled trial. *British Medical Journal*, 337(a1411).
- Hoyt, D. R., Conger, R. D., Valde, J. G., Weihs, K., and Hoyt, D. R. (1997). Psychological distress and help seeking in rural America. Technical Report 4, Sociology Department, Faculty Publications.
- Huff, D. L. (2000). Don’t misuse the Huff model in GIS. *Business Geographies*, 8(8):12.

- Hull, S. A., Jones, I. R., and Moser, K. (1997). Factors influencing the attendance rate at accident and emergency departments in East London: the contributions of practice organization, population characteristics and distance. *Journal of Health Services Research & Policy*, 2(1):6.
- Jaine, R., Baker, M., and Venugopal, K. (2008). Epidemiology of acute rheumatic fever in New Zealand 1996-2005. *Journal of Paediatrics and Child Health*, 44(10):564–571.
- Jansen, P., Bacal, K., and Buetow, S. (2011). A comparison of Maori and non-Maori experiences of general practice. *The New Zealand Medical Journal*, 124(1330):24–29.
- Johnson, J. L., Moser, L., and Garwood, C. L. (2013). Health literacy: A primer for pharmacists. *American Journal of Health-System Pharmacy*, 70(11):949–955.
- Jones, A. P., Bentham, G., Harrison, B. D., Jarvis, D., Badminton, R. M., and Wareham, N. J. (1998). Accessibility and health service utilization for asthma in Norfolk, England. *Journal of Public Health Medicine*, 20(3):312–7.
- Jones, A. P., Bentham, G., and Horwell, C. (1999). Health service accessibility and deaths from asthma. *International Journal of Epidemiology*, 28:101–105.
- Jordan, H., Roderick, P., Martin, D., and Barnett, S. (2004). Distance, rurality and the need for care: access to health services in South West England. *International Journal of Health Geographics*, 3(1):21.
- Joseph, A. E. and Bantock, P. R. (1982). Measuring potential physical accessibility to general practitioners in rural areas: a method and case study. *Social Science & Medicine*, 16:85–90.

- Kalogirou, S. and Foley, R. (2006). Health , Place and Hanly : modelling accessibility to hospitals in Ireland. *Irish Geography*, 39(1):52–68.
- Lauder, C., Skelly, C., and Brabyn, L. (2001). Developing and validating a road travel time network for cost path analysis. In *The 13th Annual Colloquium of the Spatial Information Research Centre*.
- Lawrence, P. (1990). Hanging on the Metaphone: Computer Language Magazine. *Computer Language Magazine*, 7(12).
- Letford, K. and Ashton, T. (2010). Integrated Family Health Centres. *Health Policy Monitor*, pages 1–7.
- Lin, G., Allan, D. E., and Penning, M. J. (2002). Examining distance effects on hospitalizations using GIS: a study of three health regions in British Columbia, Canada. *Environment and Planning A*, 34(11):2037–2053.
- Love, D. and Lindquist, P. (1995). The geographical accessibility of hospitals to the aged: a geographic information systems analysis within Illinois. *Health Services Research*, 29(6):629–51.
- Lovett, A., Haynes, R., Sünnerberg, G., and Gale, S. (2002). Car travel time and accessibility by bus to general practitioner services: a study using patient registers and GIS. *Social Science & Medicine*, 55(1):97–111.
- Lowthian, J. a., Curtis, A. J., Jolley, D. J., Stoelwinder, J. U., McNeil, J. J., and Cameron, P. a. (2012). Demand at the emergency department front door: 10-year trends in presentations. *The Medical Journal of Australia*, 196(2):128–132.
- Luo, W. and Qi, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place*, 15(4):1100–7.

- Luo, W. and Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design*, 30(6):865–884.
- MacRae, J. (2006). NHI Encryption Methods. *Health Care and Informatics Review Online*.
- MacRae, J., O’Malley, C., and Brown, M. (2006). Patients consulting outside of funded practices within primary health organisations: implications for utilisation reporting. *The New Zealand Medical Journal*, 119(1233):1–11.
- Magnusson, G. (1980). The role of proximity in the use of hospital emergency departments. *Sociology of Health & Illness*, 2(2):202–14.
- Manson, J., Greenland, P., VanItallie, T., and Skerrett, P. (2004). The escalating pandemics of obesity and sedentary lifestyle: A call to action for clinicians. *Archives of Internal Medicine*, 164(3):249–258.
- Martin, D., Jordan, H., and Roderick, P. (2008). Taking the bus: incorporating public transport timetable data into health care accessibility modelling. *Environment and Planning A*, 40(10):2510–2525.
- Martin, D., Roderick, P., Diamond, I., Clements, S., and Stone, N. (1998). Geographical aspects of the uptake of renal replacement therapy in England. *International Journal of Population Geography*, 4(3):227–242.
- Martin, S. and Smith, P. (1996). Explaining variations in inpatient length of stay in the National Health Service. *Journal of Health Economics*, 15(3).
- Masso, M., Bezzina, A. J., Siminski, P., Middleton, R., and Eagar, K. (2007). Why patients attend emergency departments for conditions

- potentially appropriate for primary care: reasons given by patients and clinicians differ. *Emergency Medicine Australasia*, 19(4):333–40.
- McGrail, M. R. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *International Journal of Health Geographics*, 11(1):50.
- McGrail, M. R. and Humphreys, J. S. (2009). The index of rural access: an innovative integrated approach for measuring primary care access. *BMC Health Services Research*, 9:124.
- McIntyre, D., Thiede, M., Dahlgren, G., and Whitehead, M. (2006). What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? *Social Science & Medicine*, 62(4):858–865.
- McKee, C. M., Gleadhill, D. N., and Watson, J. D. (1990). Accident and emergency attendance rates: variation among patients from different general practices. *The British Journal of General Practice*, 40(333):150–3.
- McNaughton, S. P., Collins, C. D., Kripalani, S., Rothman, R., Self, W. H., Jenkins, C., Miller, K., Arbogast, P., Naftilan, A., Dittus, R. S., and Storrow, A. B. (2013). Low Numeracy Is Associated With Increased Odds of 30-Day Emergency Department or Hospital Recidivism for Patients With Acute Heart Failure. *Circulation: Heart Failure*, 6:40–46.
- Mendelson, D. N. and Schwartz, W. B. (1993). The effects of aging and population growth on health care costs. *Health Affairs*, 12(1):119–125.
- Messina, J. P., Shortridge, A. M., Groop, R. E., Varnakovidia, P., and Finn, M. J. (2006). Evaluating Michigan’s community hospital ac-

cess: spatial methods for decision support. *International Journal of Health Geographics*, 5:42.

MidCentral District Health Board (2012). National Health Index Number. <http://www.midcentralthb.govt.nz/PatientsandVisitors/GeneralInformation/Pages/NHI.aspx>. [Online. Last accessed: 18/3/2014].

Ministry of Education (2013). Tertiary Participation Statistics. http://www.educationcounts.govt.nz/statistics/tertiary_education/participation. [Online. Last accessed: 22 December 2013].

Ministry of Health (2012). Hospital Events 2008/09 and 2009/10. Technical report, Ministry of Health, Wellington.

Ministry of Transport (2002). New Zealand transport strategy. Technical report, Ministry of Transport, Wellington.

Monnet, E., Collin-Naudet, E., Bresson-Hadni, S., Minello, A., Di Martino, V., Carel, D., Joost, V., Gagnaire, A., Evrard, P., and Obert-Clerc, B. (2006). Place of residence and distance to medical care influence the diagnosis of hepatitis C: a population-based study. *Journal of Hepatology*, 44(3):49–506.

Monnet, E., Ramée, C., Minello, A., Jooste, V., Carel, D., and Di Martino, V. (2008). Socioeconomic context, distance to primary care and detection of hepatitis C: a French population-based study. *Social Science & Medicine*, 66(5):1046–56.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, 38(8):4.

Moutier, B. C. Y., Bazzo, D. E. J., and Norcross, W. A. (2011). Approaching the Issue of the Aging Physician Population. *Journal of Medical Regulation*, 99(1):10–8.

- Nemet, G. F. and Bailey, A. J. (2000). Distance and health care utilization among the rural elderly. *Social Science & Medicine*, 50(9):1197–208.
- O'Reilly, D., Stevenson, M., McCay, C., and Jamison, J. (2001). General practice out-of-hours service, variations in use and equality in access to a doctor: a cross-sectional study. *The British Journal of General Practice*, 51(469):625–9.
- Parchman, M. and Culler, S. (1994). Primary care physicians and avoidable hospitalizations. *Journal of Family Practice*, 39(2):123–8.
- Parker, E. B. and Campbell, J. L. (1998). Measuring access to primary medical care: some examples of the use of geographical information systems. *Health & Place*, 4(2):183–93.
- Pearce, J., Witten, K., and Bartie, P. (2006). Neighbourhoods and health: a GIS approach to measuring community resource accessibility. *Journal of Epidemiology & Community Health*, 60(5):389–95.
- Pearce, J., Witten, K., Hiscock, R., and Blakely, T. (2007). Are socially disadvantaged neighbourhoods deprived of health-related community resources? *International Journal of Epidemiology*, 36(2):348–55.
- Phibbs, C. S. and Luft, H. S. (1995). Correlation of Travel Time on Roads versus Straight Line Distance. *Medical Care Research and Review*, 52(4):532–542.
- Purdy, S., Griffin, T., Salisbury, C., and Sharp, D. (2009). Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public health*, 123(2):169–73.
- Raymond, E. S. (1998). The cathedral and the bazaar. Technical Report 2.

- Reeder, A. and Trevena, J. (2003). Adult's perceptions of the causes and primary prevention of common fatal cancers in New Zealand. *The New Zealand Medical Journal*, 116(1182):1–10.
- Richardson, S., Ardagh, M., and Hider, P. (2006). New Zealand health professionals do not agree about what defines appropriate attendance at an emergency department. *The New Zealand Medical Journal*, 119(1232):U1933.
- Rizza, P., Bianco, A., Pavia, M., and Angelillo, I. F. (2007). Preventable hospitalization and access to primary health care in an area of Southern Italy. *BMC Health Services Research*, 7:134.
- Rosano, A., Loha, C. A., Falvo, R., van der Zee, J., Ricciardi, W., Guasticchi, G., and de Belvis, A. G. (2013). The relationship between avoidable hospitalization and accessibility to primary care: a systematic review. *European Journal of Public Health*, 23(3):356–60.
- Ruscoe, C. and Haran, C. (2013). Implementation of healthcare services for children under six in New Zealand. Technical report, Child Poverty Action Group.
- Sadalla, E. K. and Magel, S. G. (1980). The Perception of Traversed Distance. *Environment and Behavior*, 12(1):65–79.
- Salmond, C., Crampton, P., and Atkinson, J. (2007). NZDep2006 Index of Deprivation. Technical Report August, University of Otago, Wellington.
- Saltman, R. B. and Figueras, J. (1998). Analyzing the evidence on European health care reforms. *Health Affairs*, 17(2):85–108.
- Sauerzapf, V. a., Jones, a. P., Haynes, R., Crawford, S. M., and Forman, D. (2008). Travel time to radiotherapy and uptake of breast-conserving surgery for early stage cancer in Northern England. *Health & Place*, 14(3):424–33.

- Schuurman, N., Fiedler, R. S., Grzybowski, S. C. W., and Grund, D. (2006). Defining rational hospital catchments for non-urban areas based on travel-time. *International Journal of Health Geographics*, 5:43.
- Sheerin, I., Allen, G., Henare, M., and Craig, K. (2006). Avoidable hospitalisations: potential for primary and public health initiatives in Canterbury, New Zealand. *The New Zealand Medical Journal*, 119(1236):U2029.
- Siedner, M., Lankowski, A., Tasi, A., Muzoora, C., Martin, J., Hunt, P., Haberer, J., and Bangsberg, D. (2013). GP-measured distance to clinic, but not self-reported transportation factors, are associated with missed HIV clinic visits in rural Uganda. *AIDS*, 27(9):1503–8.
- Siminski, P., Bezzina, A. J., Lago, L. P., and Eagar, K. (2008). Primary care presentations at emergency departments: rates and reasons by age and sex. *Australian health review : a publication of the Australian Hospital Association*, 32(4):700–9.
- Statistics New Zealand (2001). Census of population and dwellings: families and households. Technical report, Statistics New Zealand, Wellington.
- Statistics New Zealand (2013). Meshblocks. http://www.stats.govt.nz/surveys_and_methods/methods/classifications-and-standards/classification-related-stats-standards/meshblock/definition.aspx. [Online. Last accessed: 3 April 2013].
- Stock, R. (1983). Distance and the utilization of health facilities in rural Nigeria. *Social Science & Medicine*, 17(9):563–70.
- Suhrchke, M., Nugent, R., Stuckler, D., and Lorenzo, R. (2006). Chronic disease: an economic perspective. Technical Report November, Oxford Health Alliance, London.

- Teach, S. J., Guagliardo, M. F., Crain, E. F., McCarter, R. J., Quint, D. M., Shao, C., and Joseph, J. G. (2006). Spatial accessibility of primary care pediatric services in an urban environment: association with asthma management and outcome. *Pediatrics*, 117(4 Pt 2):S78–85.
- Thompson, S. (1998). Today’s actions tomorrow’s landscapes. In *Anniversary Conference of the New Zealand Institute of Landscape Architects.*, Wellington, New Zealand.
- Timmins, C. L. (2002). The impact of language barriers on the health care of Latinos in the United States: a review of the literature and guidelines for practice. *Journal of Midwifery & Women’s Health*, 47(2):80–96.
- Tiwari, C. and Rushton, G. (2005). Using spatially adaptive filters to map late stage colorectal cancer incidence in iowa. In *Developments in Spatial Data Handling, Proceedings of the 11th International Symposium on Spatial Data Handling.*, pages 665–676, Berlin. Springer.
- Turnbull, J., Martin, D., Lattimer, V., Pope, C., and Culliford, D. (2008). Does distance matter? Geographical variation in GP out-of-hours service use: an observational study. *The British Journal of General Practice*, 58(552):471–7.
- Wang, F. and Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health & Place*, 11(2):131–46.
- Webb, T. L., Chang, B. P. I., and Benn, Y. (2013). The Ostrich Problem: Motivated Avoidance or Rejection of Information About Goal Progress. *Social and Personality Psychology Compass*, 7(11):794–807.
- Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. *Regional Science and Urban Economics*, 6:357–379.

- Wilson, H. (2005). Co-locating primary care facilities within emergency departments: brilliant innovation or unwelcome intervention into clinical care? *The New Zealand Medical Journal*, 118(1221):U1633.
- Wilson, N. (2010). Rheumatic heart disease in indigenous populations—New Zealand experience. *Heart, Lung & Circulation*, 19(5-6):282–8.
- Witten, K., Exeter, D., and Field, A. (2003). The Quality of Urban Environments: Mapping Variation in Access to Community Resources. *Urban Studies*, 40(1):161–177.
- Yang, D., Goerge, R., and R, M. (2006). Comparing GIS-based methods of measuring spatial accessibility to health services. *Journal of Medical Systems*, 30(1):23–32.

Appendices

Appendix A

Interpolation Methodology

Figures A.1 and A.2 show the various dimensions that were used and calculated in order to achieve interpolated values for each road network node (r) in the horizontal and vertical planes respectively.

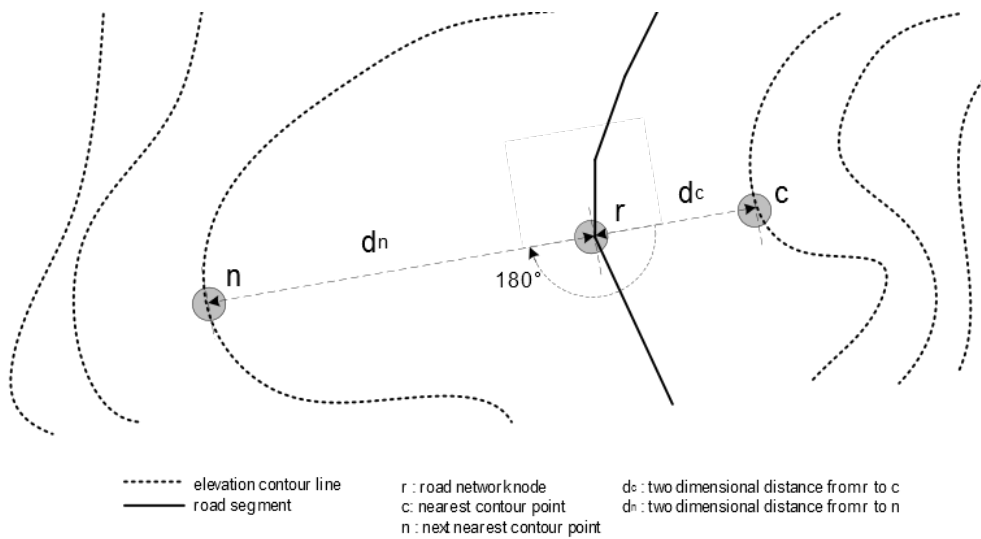


Figure A.1: Diagram showing interpolation dimensions on a horizontal plane (birds-eye view)

The purpose of the algorithm is to determine the road network node elevation, based on interpolated values of contour reference points. The algorithm locates the closest contour point (c) to each road network node on the horizontal plane. It then locates the closest contour point in a direction 180 degrees on the horizontal to the bearing formed between the road network node and the closest contour point, and this

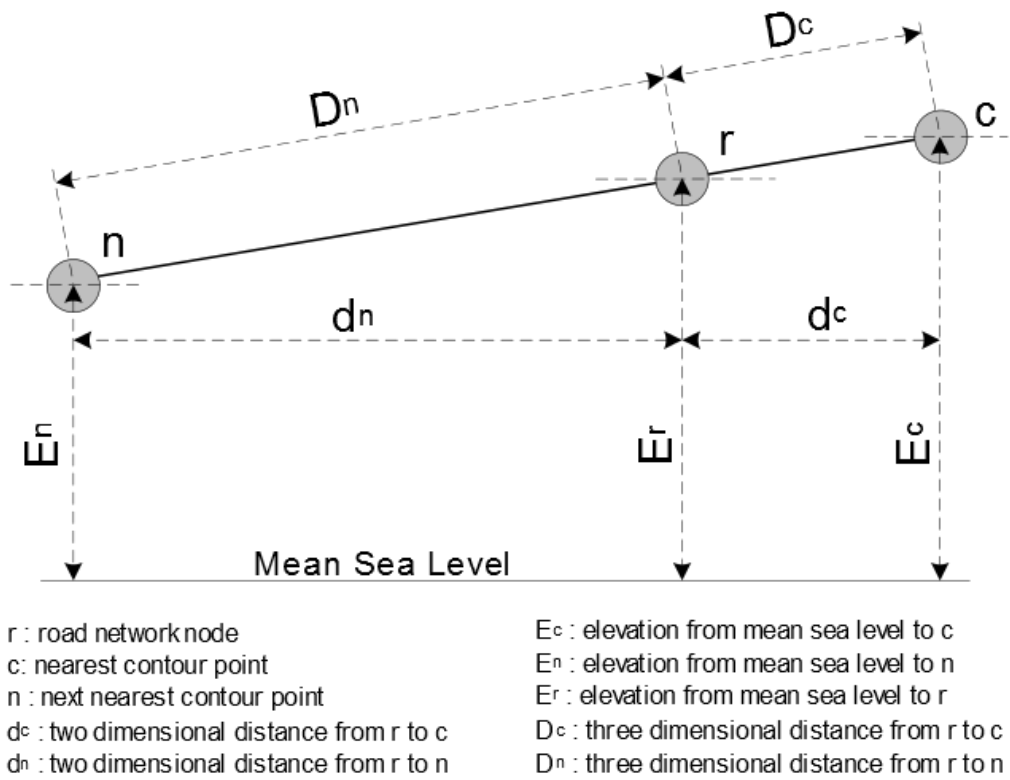


Figure A.2: Diagram showing interpolation dimensions on a vertical plane (cross-section view)

is referred to as the next closest contour point (n). The elevations of these two contour points (E_c and E_n) are known from the contour data, and the two dimensional Euclidean distances between the road network node and these contour points (d_c and d_n) are calculated. Trigonometry provides the basis for determining the three dimensional distance (D_c and D_n) between the road network node and the contour points.

The road network node's elevation is calculated by determining the elevation of the linear gradient between the two contour points on which the node lies. It can be represented as follows:

$$E_r = E_c \left(\frac{D_n}{D_n + D_c} \right) + E_n \left(\frac{D_c}{D_n + D_c} \right)$$

where :

$$D_c = \frac{d_c}{\cos \left(\tan^{-1} \left(\frac{|E_c - E_n|}{d_n + d_c} \right) \right)}$$

and :

$$D_n = \frac{d_n}{\cos \left(\tan^{-1} \left(\frac{|E_c - E_n|}{d_n + d_c} \right) \right)}$$

Appendix B

Algorithm Implementation Details

1. MacRae Road Network Travel Time Model
PostGIS
http://bit.ly/macrae_travel_time_model
2. Brabyn and Skelly Road Network Travel Time Model
PostGIS
http://bit.ly/macrae_bsttm_postgis
3. Path Sinuosity Calculations using Dynamic Segmentation
PostGIS
http://bit.ly/macrae_sinuosity_index
http://bit.ly/macrae_sinuosity_postgis
4. Splitting NZ Addresses into Address Components
Regular Expression
http://bit.ly/macrae_address_components
5. Aggregating GPS Points to Line Geometries
PostGIS
http://bit.ly/macrae_aggregate_gps
6. Calculating Turning Angle of Geometry
PostGIS
http://bit.ly/macrae_turning_angle